

**GENETIC AND BIOCHEMICAL ANALYSES OF HYPOTHETICAL
PROTEIN 1: AN INTERACTING PARTNER OF CIKA IN
Synechococcus elongatus PCC 7942**

A Thesis

by

HAITAO GUO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

May 2007

Major Subject: Microbiology

**GENETIC AND BIOCHEMICAL ANALYSES OF HYPOTHETICAL
PROTEIN 1: AN INTERACTING PARTNER OF CIKA IN
Synechococcus elongatus PCC 7942**

A Thesis

by

HAITAO GUO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

| | |
|---------------------|-----------------------|
| Chair of Committee, | Susan S. Golden |
| Committee Members, | Deborah Bell-Pedersen |
| | Andy Liwang |
| | Jin Xiong |
| Head of Department, | Vincent M. Cassone |

May 2007

Major Subject: Microbiology

ABSTRACT

Genetic and Biochemical Analyses of Hypothetical Protein 1: An Interacting Partner of CikA in *Synechococcus elongatus* PCC 7942. (May 2007)

Haitao Guo, B.S., Central China Normal University;
M.S., Institute of Hydrobiology, Chinese Academy of Sciences
Chair of Advisory Committee: Dr. Susan S. Golden

Synechococcus elongatus PCC 7942 is a model organism used to study the circadian rhythm, a process that is driven by an endogenous biological clock that can be modulated by external cues such as light and temperature. Some proteins have been identified that are involved in circadian signal transduction in *S. elongatus*. Of them, KaiA, KaiB and KaiC comprise the central oscillator components, which are essential for internal timekeeping. SasA is an important protein in the output pathway, which passes the information from central oscillator to downstream components, and thus controls metabolic and behavioral processes. CikA is a major component in the input pathway, which maintains synchrony of the oscillator with the environment. CikA is an unusual phytochrome-like histidine protein kinase. It has a *pseudo* receiver domain which can not accept a phosphoryl group. CikA is thought to be located at the poles of the cell through interaction between PsR and some protein or protein complex that is also localized at the poles. One of the potential CikA-interacting proteins identified through a yeast two hybrid screen is called hypothetical protein 1. It specifically recognizes a PsR bait in a yeast two hybrid assay. A bioinformatics analysis showed that there are predicted signal peptide and transmembrane domains at the N-terminal and a cytochrome C homolog domain at the C-terminal of Hyp1. Elucidating the features and function of Hyp1 will provide us with useful information to understand the function and working mechanism of CikA, and therefore will help us to clarify the signal transduction in the clock. In this research, I used genetic, cell biological and biochemical approaches to study the features and function of this newly identified clock component Hyp1.

To confirm the interaction between PsR and Hyp1 and complement the yeast two hybrid data, I truncated Hyp1 (Thyp1) and purified soluble Thyp1. At the same time, I obtained purified PsR. I tried to copurify the PsR and 6-histidine-tagged Hyp1 on a nickel

affinity column. However, PsR non-specifically bound to the column, which eliminated the utility of this approach to study their interaction.

In addition to using a biochemical approach to study Hyp1, I constructed three *hyp1* overexpression alleles for genetic analysis and two *hyp1-yfp* overexpression fusion alleles for subcellular localization studies. All of them will help us to understand the features and function of Hyp1.

DEDICATION

To my father and mother

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Susan S. Golden, for her continuous support of my research. She was always there to meet and talk about ideas, to proofread and edit my research proposal and thesis. Without her help, I could not have completed this thesis. She showed her respect for my choice and gave me a lot of encouragement to pursue my interests, even when the interests went beyond science.

Besides my advisor, I would like to thank the rest of my committee members, Dr. Deborah Bell-Pedersen, Dr. Andy Liwang and Dr. Jin Xiong, for their guidance and insightful comments and hard questions. They taught me how to think scientifically and critically.

I also want to say “thank you” to the following people at Texas A&M University: Dr. James W. Golden, who challenged me with a lot of good questions in lab meetings; Dr. Deborah Siegele, who invited our international students to celebrate the first Turkey Day in America; Dr. Kay Goldman, who dedicated her precious time to solve any problems concerning graduate students.

I am also greatly indebted to many labmates and friends for their help in the past three years. They made my time at Texas A&M University a great experience.

Last, but not least, I thank my father and mother, Houyun Guo and Shaoyue Zhang, for their great love.

NOMENCLATURE

| | |
|------------|---|
| 6XHis | Six adjacent histidine residues |
| C-terminus | Carboxy-terminus |
| C-Hyp1 | C terminus of Hyp1 |
| CikA | Circadian input kinase |
| HPK | Histidine protein kinase |
| Hyp1 | Hypothetical protein 1 |
| IPTG | Isopropyl-beta-D- thiogalactopyranoside |
| kD | Kilodalton |
| LdpA | Light dependent period protein |
| N-Hyp1 | N terminus of Hyp1 |
| Ni-NTA | Nickel-nitrilotriacetic acid |
| NS | Neutral site |
| N-terminus | Amino terminus |
| OD | Optical density |
| ORF | Open reading frame |
| PCR | Polymerase chain reaction |
| PsR | <i>Pseudo</i> -receiver |
| RR | Response regulator |
| SasA | <i>Synechococcus</i> adaptive sensor |
| SDS-PAGE | Sodium dodecyl sulfate - polyacrylamide gel electrophoresis |
| Sm | Streptomycin |
| Sp | Spectinomycin |
| Thyp1 | Truncated Hypothetical protein 1 |

TABLE OF CONTENTS

| | Page |
|--|------|
| ABSTRACT | iii |
| DEDICATION | v |
| ACKNOWLEDGEMENTS | vi |
| NOMENCLATURE | vii |
| TABLE OF CONTENTS | viii |
| LIST OF FIGURES | x |
| LIST OF TABLES | xi |
| CHAPTER | |
| I INTRODUCTION | 1 |
| Circadian rhythms | 1 |
| The cyanobacterial clock | 2 |
| Central oscillators | 4 |
| Output pathway | 6 |
| Input pathway | 7 |
| Current model of the cyanobacterial clock | 10 |
| Interaction partners of CikA from a yeast two-hybrid assay | 12 |
| Bioinformatics and genetics information regarding Hyp1 | 13 |
| II USING A PULL-DOWN ASSAY TO CONFIRM THE INTERACTION BETWEEN PsR AND HYP1 <i>IN VITRO</i> | 15 |
| Introduction | 15 |
| Materials and methods | 16 |
| Bacterial strains and culture conditions | 16 |
| DNA manipulations and construction of overexpression plasmids | 17 |
| Quick-change mutagenesis of the PsR construct | 17 |
| Protein overexpression and analysis | 18 |
| Pull-down assay | 19 |
| Immunoblot analyses | 20 |
| Results | 21 |
| Bioinformatics analysis of Hyp1 | 21 |
| Optimizing truncated Hyp1 protein overexpression | 22 |
| Overexpression of PsR without a 6XHis tag | 24 |
| Co-purification of PsR and Thyp1 | 24 |
| Discussion | 29 |

| CHAPTER | Page |
|--|------|
| III CONSTRUCTION OF THREE HIS-TAGGED HYP1 VARIANTS AND TWO HYP1-YELLOW FLUORESCENT PROTEIN FUSIONS FOR GENETIC AND SUBCELLULAR LOCALIZATION ANALYSES OF HYP1 | 31 |
| Introduction..... | 31 |
| Material and methods..... | 32 |
| Bacterial strains and plasmids..... | 32 |
| Constructing three C-terminal His-tagged Hyp1 overexpression Variants | 32 |
| Constructing full-length <i>hyp1-yfp</i> and truncated <i>hyp1-yfp</i> overexpression alleles | 34 |
| Visualization of intracellular localization of YFP | 35 |
| Results and discussion | 35 |
| Constructing three C-terminal His-tagged Hyp1 overexpression Variants | 35 |
| Construction and overexpression of full-length <i>hyp1-yfp</i> and truncated <i>hyp1-yfp</i> fusion alleles | 37 |
| IV SUMMARY AND CONCLUSIONS..... | 40 |
| REFERENCES | 41 |
| APPENDIX A A SUMMARY OF USING A YEAST TWO HYBRID SYSTEM TO IDENTIFY PROTEINS THAT INTERACT WITH KAIB..... | 44 |
| VITA..... | 49 |

LIST OF FIGURES

| FIGURE | Page |
|--|------|
| 1 Characteristics of the <i>Synechococcus elongatus</i> PCC 7942 circadian rhythm | 3 |
| 2 The circadian clock in cyanobacteria..... | 4 |
| 3 The transcription-translation feedback loop and the post-translational oscillatory feedback loop | 5 |
| 4 Three domains of CikA: CikA includes GAF, histidine protein kinase (H and ATP) and receiver-like (PsR) domains | 8 |
| 5 The inferred structure for CikA: A and B | 10 |
| 6 A current model of the cyanobacterial clock | 11 |
| 7 Yeast two-hybrid assay | 12 |
| 8 Protein affinity chromatography | 15 |
| 9 Bioinformatics analysis of Hyp1 | 21 |
| 10 Overexpression of TRX-HIS-Thyp1..... | 22 |
| 11 Thyp1 protein expression was optimized to get maximum soluble protein | 23 |
| 12 The overexpression of PsR after genetically removing the 6XHis..... | 25 |
| 13 PsR was detected in the eluate independent of presence of His-Thyp1 | 26 |
| 14 Pull-down assays with different washing stringencies | 28 |
| 15 Schematic representation of each construct..... | 39 |
| 16 Flowchart of yeast two-hybrid assay | 45 |
| 17 The gene structure of the bacterioferritin comigratory protein locus | 47 |

LIST OF TABLES

| TABLE | Page |
|---|------|
| 1 Plasmids used in this study | 33 |
| 2 Primers used for cloning | 34 |
| 3 The sequencing and retest results of 17 positive clones | 46 |

CHAPTER I

INTRODUCTION

The research presented in this thesis used genetic and biochemical approaches to understand the features and function of a newly identified clock component termed Hypothetical protein 1.

Circadian rhythms

The rotation of the earth generates an alternation of day and night. To exploit predictable environmental changes that this daily cycle generates, most living organisms have evolved circadian rhythms. A circadian rhythm is a cyclic behavior that is driven by an endogenous biological clock that can be modulated by external cues such as light and temperature (2, 3). Period, phase, and amplitude are three terms that are frequently used to describe a rhythm. Period is the time after which a defined point of an oscillation recurs, phase is the instantaneous state of an oscillation within a period (for example, the timing of a peak relative to a reference point), and amplitude is deviation from the mean of an oscillation to its maximum or minimum. Three criteria must be satisfied to determine that a rhythm is circadian: (1) The rhythm persists in constant conditions (for example constant dark) with a period of about 24 hours, which means that a circadian rhythm is a product of an endogenous activity of the organism rather than driven by an external stimulus. Even without an environmental cycle, it can freely run. (2) The phase of rhythm can be reset by exposure to an environmental cue, such as a light or dark pulse. Though the environmental stimulus will not determine the cycle itself, it can change the relative phase. The phase of the rhythm must be able to be synchronized to local time to avoid being inconsistent with its surroundings. (3) The rhythm proceeds at the same rate

This thesis follows the style of *Proceedings of the National Academy of Sciences of the United States of America*.

within a range of temperatures that are normal for the organism's natural habitat. This property, called temperature compensation, ensures that the biological clock can maintain a robust cycling within the organism's physiological range in a natural setting.

Circadian rhythms have been found in most organisms examined so far, from a unicellular cyanobacterium and filamentous fungi to fruit flies and mammals (4) (Bell-Pedersen, *et al.* 2005). Cyanobacteria are so far the only group of prokaryotes and the simplest organisms demonstrated to possess a *bona fide* circadian clock (4) (Bell-Pedersen, *et al.* 2005).

The cyanobacterial clock

Synechococcus elongatus PCC 7942 is a unicellular freshwater cyanobacterium that is obligately photoautotrophic. It has several features that favor genetic manipulation (5, 6): (1) It is naturally transformable and homologous recombination can occur with high efficiency (7). (2) It has a relatively small genome size (~2.7 Mb) and its complete genome sequence has been released (http://genome.ornl.gov/microbial/syn_PCC7942). It has become one of the model organisms in circadian rhythms research and the only developed model for a prokaryotic clock (4). Using a promoter-less bacterial luciferase gene set (*luxAB*) from the marine bacterium *Vibrio harveyi* as a reporter, we can automatically monitor the bioluminescence that results from circadian expression of promoters in *S. elongatus* (8, 9). Period, phase, and amplitude of circadian rhythms can be easily discerned from the bioluminescence rhythms (Figure 1).

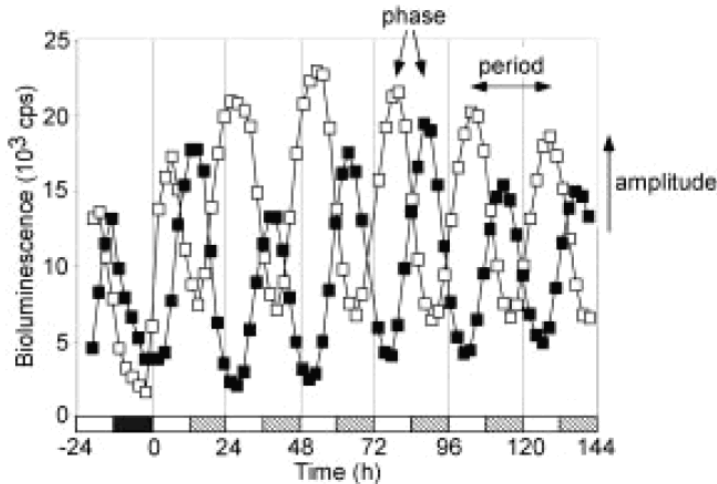


Fig. 1. Characteristics of the *Synechococcus elongatus* PCC 7942 circadian rhythm. Negative time denotes time during light–dark (LD) cycles that synchronize the cells. The black bar indicates time in darkness during the LD cycle; hatched bars indicate “subjective” dark during constant light (LL) conditions. Properties of the curve that are typically measured are the period (the time between occurrence of peaks or troughs), phase (the relative positioning of the curve with respect to time entering LL), and amplitude (the expression level measured from the midline of the curve to either the peak or trough) of the rhythm. Time points from a *PkaiBC::luxAB* reporter (AMC462; closed squares) show a class 1 phase rhythm, peaking at the light to dark transition, or subjective dusk. *PpurF::luxAB* (AMC408; open squares) represent a class 2 phase reporter that peaks at subjective dawn, 12 h out of phase from class 1 rhythms in constant conditions. (Figure and legend used, with permission, from S.S. Golden)

Since the genes that comprise the circadian oscillator were identified in *S. elongatus* in 1998 (10), a multi-lab effort to understand the circadian mechanism in this model organism has made remarkable progress.

Circadian clock systems consist of three major parts: an input pathway (which senses environmental cues, such as light and temperature, and passes the signal to the central oscillator), a central oscillator (which is the timekeeper itself and creates a near 24-hour circuit), and an output pathway (which gets the information from the central oscillator and regulates downstream gene expression, physiology or behavior) (Figure 2) (5).

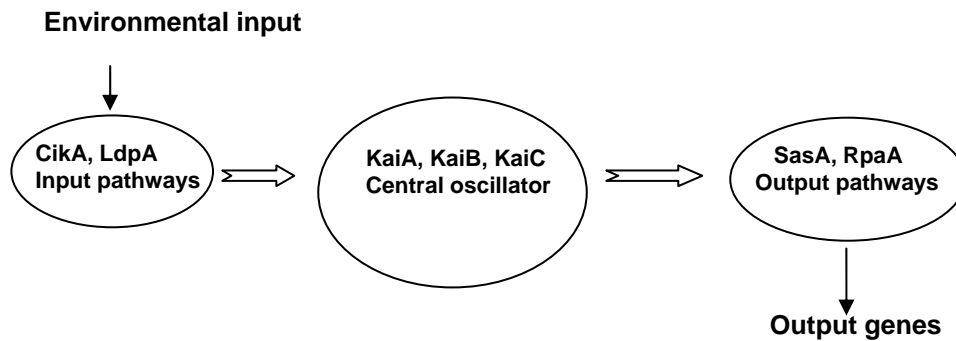


Fig. 2. The circadian clock in cyanobacteria. The circadian clock in cyanobacteria is composed of a central oscillator, input pathways, and output pathways. The key components of each of these are shown in the figure.

Central oscillators

The *kaiA*, *kaiB*, and *kaiC* genes were originally identified in 1998 with a chemical mutagenesis approach (10). They comprise a *kai* gene cluster in the genome (Figure 3). Deleting any one of *kai* genes will render the clock arrhythmic. Thus, the *kai* genes are thought to be central oscillator components (10).

kaiB and *kaiC* are in an operon downstream of *kaiA* gene, and *kaiA* has its own promoter (10). KaiA, KaiB, and KaiC are not homologous to key clock components from major groups of eukaryotic organisms, but they have similar auto-regulatory properties (10). The KaiC protein is a repressor for *kaiBC* expression: overexpression of KaiC will inhibit its own expression. Conversely, KaiA is a positive regulator of *kaiBC* and overexpression of KaiA will increase *kaiBC* transcription. Taken all together, KaiA, KaiB, and KaiC form a negative transcription-translation feedback loop. (Figure 3)

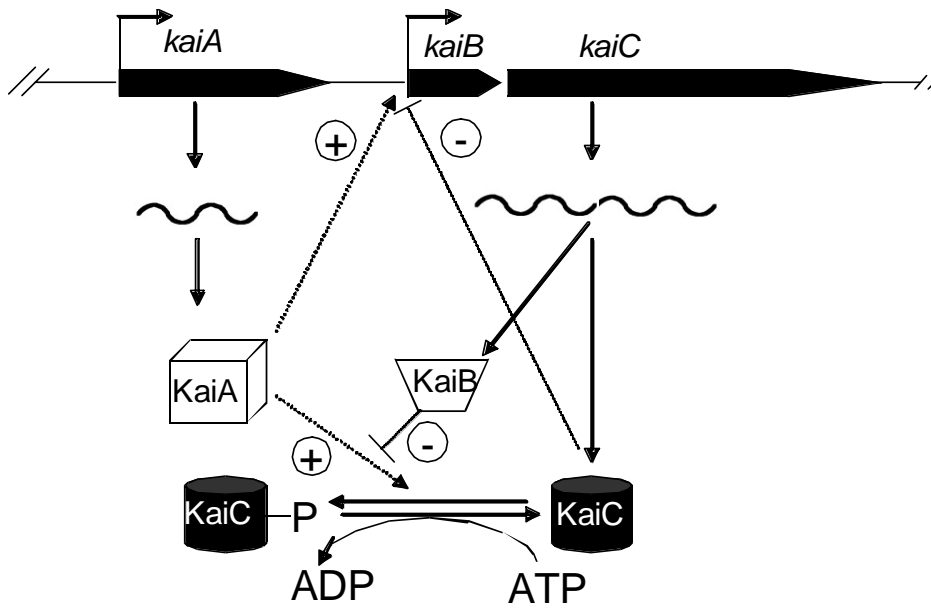


Fig. 3. The transcription-translation feedback loop and the post-translational oscillatory feedback loop. *kaiA*, *kaiB*, and *kaiC* comprise a *kai* gene cluster in the genome. *kaiA* has its own promoter. *kaiB* and *kaiC* form an operon. Overexpression of KaiA will stimulate *kaiBC* expression, while KaiC overexpression will inhibit *kaiBC* transcription. Together, they form a transcription-translation feedback loop. At the post-translational level, KaiC's autophosphorylation activity is enhanced by KaiA, while KaiB weakens KaiA's effect, which forms a post-translational oscillatory feedback loop.

Bioinformatics failed to detect any identifiable motifs in KaiA or KaiB, but KaiC carries recognizable ATP-binding motifs typical of the RecA/DnaB superfamily (10). The N-terminal and C-terminal domains of KaiC are similar to each other; each domain contains a Walker A motif that can bind ATP (10, 11). Like other members of this superfamily, KaiC forms homomultimers (specifically, a hexameric ring) in the presence of ATP (12, 13, 14). KaiC protein autophosphorylates when incubated with radiolabeled ATP. There are three phosphorylation sites. Two of them are shown to exist on adjacent serine and threonine residues (11, 14, 15). During the course of the circadian cycle, KaiC abundance and phosphorylation states change (11). KaiB abundance cycles in a circadian manner in synchrony with KaiC, but KaiA levels are constitutive or oscillate with a very low amplitude (16). KaiC's autophosphorylation activity is enhanced by KaiA, while KaiB antagonizes KaiA's effect (17, 18), which forms a post-translational oscillatory feedback loop (Figure 3).

A transcription/translation oscillatory (TTO) feedback oscillator is the most commonly used mechanism underlying eukaryotic circadian systems (19), and cyanobacteria were expected to use a similar TTO feedback mechanism (20). However, recent data demonstrated that the phosphorylation state of the oscillator protein KaiC plays a critical role, independent of transcription or translation. The three Kai proteins and ATP can reconstitute a circadian oscillation of the phosphorylation pattern *in vitro*, similar to that *in vivo*, that is sustained for at least three days (21). These data show that the cyanobacteria use a post-translational oscillator (22).

Although bioinformatics analyses of KaiA offered no hints to its function, the structural information did offer some clues. KaiA forms a dimer in solution through interactions in its C-terminus (23). This C-terminal region can bind a small peptide from KaiC, and is alone responsible for the stimulation of KaiC autophosphorylation. The N-terminal region of KaiA resembles, structurally, the *pseudo-receiver* domain of CikA, which is a key player in the input pathway. The current hypothesis for the *function* of the N-terminus of KaiA is that it accepts the signal from the input pathway and then passes it to the Kai oscillator through regulating the autophosphorylation status of KaiC (24). KaiB is the clock protein whose function we know the least about. In some way it blocks the stimulation of KaiC phosphorylation by KaiA. The X-ray crystal structure of KaiB reveals that it is comprised of four subunits that are organized into two asymmetric dimers (25).

Output pathway

The *Synechococcus* adaptive sensor (SasA) protein is a sensory histidine kinase that directly associates with KaiC (26). Disruption of the *sasA* gene will result in three phenotypes: (1) dramatically reducing amplitude of the *kai* expression rhythms; (2) shortening the period of the *kai* expression rhythms; and (3) attenuating circadian expression patterns of all tested genes, to the point that most appear arrhythmic (26). Therefore, SasA plays an important role in output pathway signal transduction.

At the N-terminus of SasA, there is a sensory domain that shares sequence similarity with KaiB, but no interaction between KaiB and SasA was detected *in vitro* or *in vivo*. However, this domain was found to interact with KaiC in a yeast two-hybrid

system (26). SasA can be autophosphorylated at a histidyl residue. This autophosphorylation is crucial to its function as one amino acid substitution at the phosphorylation site will phenocopy the *sasA* null (26). KaiC will enhance SasA autophosphorylation, while KaiA and KaiB lessen the effect of KaiC on SasA (27).

Nearly all the genes expression in the *S. elongatus* genome are under circadian control, which indicates that there is a global regulatory mechanism. One puzzle is how the Kai proteins drive circadian rhythms in a genome-wide scale? Recent research progress in two groups gives some answers: Dr. Stanly B. Williams' research group observed a rhythm in the compaction/decompaction of the cyanobacterial chromosome as a function of circadian time (27). Presumably, the degree of compaction would then control access of the transcription machinery to particular promoter elements (27). In this way, the circadian clock in *S. elongatus* might control global gene expression.

At nearly at the same time, the research group in Dr. Hideo Iwasaki's lab reported a presumed DNA-binding response regulator, RpaA, as a SasA partner that is involved in the Kai-based circadian system (28). RpaA is required for genome-wide transcription rhythms. Their results strongly suggest that SasA-RpaA form a two-component system which may be the primary output pathway linking the Kai protein oscillator to circadian transcription (28). It is still not clear how chromosome compaction dynamics is connected to regulation by the SasA-RpaA two-component system.

Input pathway

There are two key components of the input pathway in *S. elongatus*: CikA (circadian input kinase) and LdpA (light-dependent period A) (29).

The *ldpA* gene was identified from transposon-based mutagenesis in a screen for mutants that respond differently than wild type to a phase-resetting dark pulse. Additional experiments showed that an *ldpA* null mutant is fully competent to reset to an external cue, which means that *ldpA* does not affect non-parametric (phase-resetting) signals. *S. elongatus* follows Aschoff's rule, in which the circadian period of diurnal organisms is slightly shorter under higher light intensities than at lower intensities, but in the *ldpA* mutant background, these typical period variations are abrogated. In this sense, LdpA does appear to act in the input pathway, affecting parametric (period-altering) signals (30).

LdpA contains two Fe-S centers, and is redox sensitive. Affinity purification of His-LdpA from the cyanobacterium showed that CikA, KaiA, and SasA co-purify with LdpA, and that LdpA association with CikA and KaiA oscillates in a circadian fashion. The current model is that LdpA senses the redox state of the cell and then transduces the information to the Kai oscillator (29).

CikA is an unusual phytochrome-like histidine protein kinase. *cikA* mutants of *S. elongatus* fail to reset the phase of the circadian rhythm of gene expression after an environmental time cue and also exhibit reduced amplitude and shortened period of circadian oscillation. In addition, the disruption of CikA results in a cell division defect (31).

Bioinformatics results predict that CikA contains three domains: a histidine protein kinase (HPK) domain, a GAF (GAF) domain and a *pseudo*-receiver (PsR) domain (32) (Figure 4).

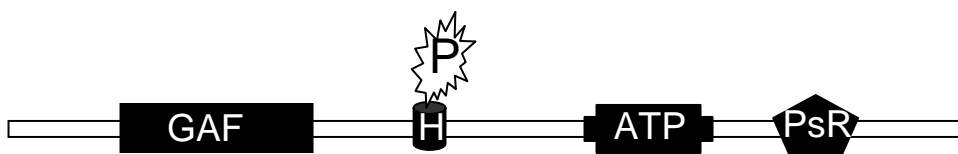


Fig. 4. Three domains of CikA: CikA includes GAF, histidine protein kinase (H and ATP), and receiver-like (PsR) domains. The GAF domain lacks a conserved cysteine residue that the bilin can attach to. The PsR domain does not have a conserved aspartyl residue to accept the phosphoryl group.

Typically, a GAF domain will bind a bilin chromophore in bacterial photoreceptors, but the CikA GAF does not contain the conserved cysteine residue to which the bilin covalently attaches. Purifying the protein from *S. elongatus* cells did not reveal an attached bilin, which makes us hypothesize that CikA does not act as a photoreceptor (33). Recent data in our lab showed that CikA might sense light indirectly by measuring the redox state of the PQ pool and then pass this information to the oscillator (34).

Autophosphorylation assays established that CikA is a *bona fide* kinase. The HPK domain can undergo autophosphorylation both *in vitro* and *in vivo*. This phosphorylation is essential for CikA function. One amino acid substitution at the phosphorylation site

results in inactivation of the protein (32, 33). HPK proteins are usually part of two-component bacterial signal transduction systems (35). The HPK protein autophosphorylates in response to a signal and then transfers that phosphoryl group to the receiver domain of a partner response regulator (RR) protein. The highly-conserved HPK domain and autokinase activity of CikA suggest that CikA is part of a bacterial two-component signal transduction system with a hypothetical partner RR, named CikR. At this time, CikR has not been identified.

The C-terminus of CikA contains a receiver-like domain, but it turned out to be a *pseudo*-receiver. PsR of CikA lacks the conserved aspartyl residue that serves as a phosphoryl acceptor in the genuine receiver domains of RRs, and cannot be phosphorylated (32). Thus, PsR is not the cognate RR for the CikA HPK domain. Dr. Xiaofan Zhang's recent results in this laboratory showed that the PsR domain has multiple functions: (1) it negatively regulates the kinase activity of HPK through intramolecular interaction; (2) it is critical for the *in vivo* localization of CikA to the poles of the cell; and (3) it is hypothesized to be involved in protein-protein interactions with other partners (33). More recently, Dr. Natalia B. Ivleva discovered that DBMIB (2, 5-dibromo-3-methyl-6-isopropyl-*p*-benzoquinone), a quinone analog, binds to the PsR domain directly and changes the structure of the PsR domain (34). It demonstrates an unrecognized ligand-binding role for the receiver fold. The presence of DBMIB has a dramatic effect on CikA stability (34).

The HPK can be divided into two subdomains, an ATP-binding domain and a histidine phosphotransfer domain (DHp), which are well conserved among HPK proteins. EnvZ is a typical HPK protein, whose structure has been partially determined (Figure 5A) (36). Based on the similarity of kinase domains and their function, the C-terminal portion of dimeric CikA is likely to have a structure similar to EnvZ (Figure 5B). The DHp domain of CikA is hypothesized to be the interaction site for CikR. It will dock and receive a phosphoryl group from the catalytic H box. The negative regulation of kinase by PsR *in vitro* and the structural similarity of PsR to RR receivers as determined by NMR data lead us to propose that the receiver-like PsR domain also interacts with the DHp domain (Figure. 5B) and blocks the interaction of CikA with CikR when CikA is not docked to its binding site. Dr. Xiaofan Zhang proposed an inferred structure and

function model for CikA based on the above hypothesis: PsR docks CikA to the cell poles through interaction with some proteins in the cell membrane. The correct localization of CikA will induce the protein to form an open structure which will activate the kinase. HPK domain is exposed and the presumed cognate RR, CikR will interact with CikA to complete the phosphor transfer (33).

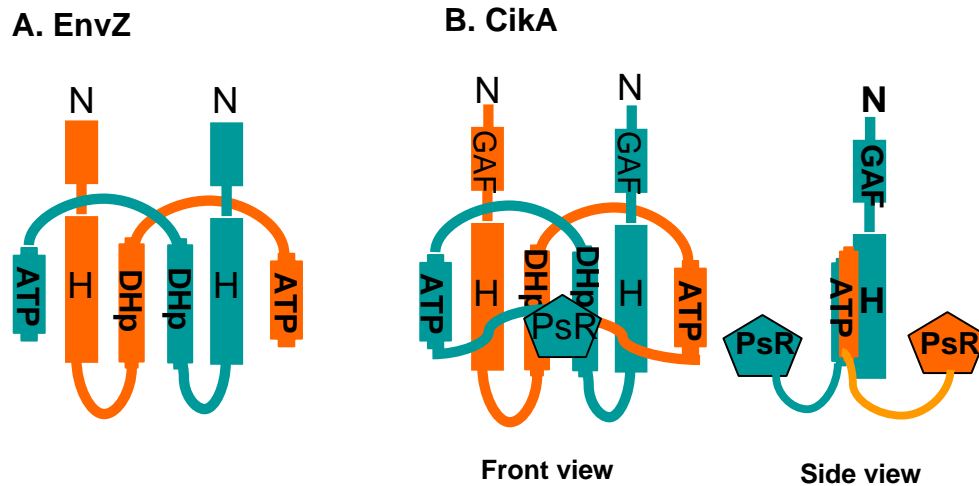


Fig. 5. The inferred structure for CikA: A and B. Structure of EnvZ (A), and inferred structure of CikA (B). For EnvZ, only the cytoplasmic catalytic region is represented. Both EnvZ and CikA are dimers. Different colors represent different monomers. (adapted from reference 33).

Current model of the cyanobacterial clock

The three Kai proteins (in various multimeric states), SasA, and likely CikA (34) and some other proteins, form a high molecular weight (mw) complex, the periodosome. The periodosome assembles maximally in the late night and disassembles in the late morning (37). Formation and dissolution of the periodosome over the course of the circadian cycle is important for the timing mechanism *in vivo*. LdpA and CikA will sense a light signal indirectly by sensing the redox status of the cell. They will transmit environmental input to the central oscillator via a hypothetical cognate response regulator, CikR, or perhaps directly through the *pseudo*-receiver N-terminus of KaiA. This stimulates the autophosphorylation of KaiC and phosphorylation of KaiC is necessary for periodosome

formation. SasA autophosphorylation is stimulated by phosphorylated KaiC. SasA then will transfer the phosphoryl group to its cognate response regulator, RpaA. Phosphorylated RpaA regulates rhythmic expression of the downstream target genes through some undefined mechanism, and thus produces overt rhythms. (Figure 6).

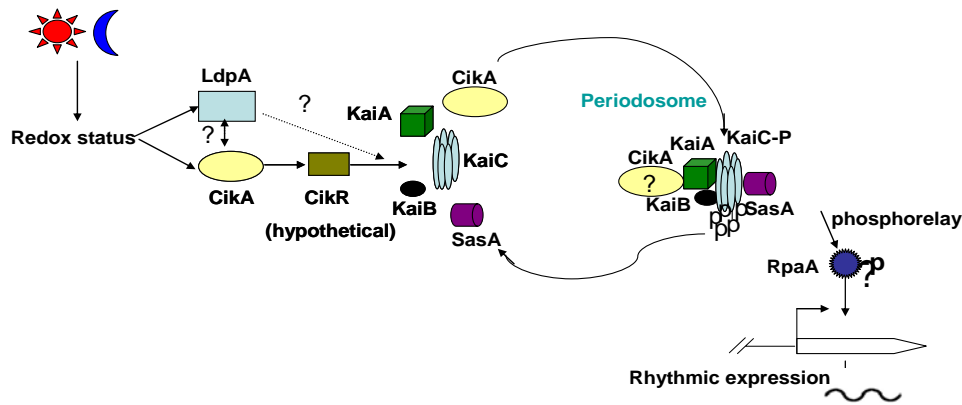


Fig. 6. A current model of the cyanobacterial clock. The three Kai proteins, SasA, and likely CikA and some other proteins, assemble a high molecular weight complex termed the periodosome. The periodosome will form and disassemble over the course of the circadian cycle, which is necessary for a fully functional clock. Environmental input is transmitted to the central oscillator via LdpA, CikA, and possibly CikR. This input stimulates the autophosphorylation of KaiC. SasA autophosphorylation is stimulated by phosphorylated KaiC. Phosphorylated SasA will initiate the phosphorelay to its cognate response regulator, RpaA. Phosphorylated RpaA activates target genes through some mechanisms that are not very well known.

Through years of efforts by different labs, we now have a relatively complete picture of the cyanobacterial circadian clock system. However, there are still gaps that need to be filled in. One of our lab's focuses is on revealing the proteins that are involved in the input pathway of the clock system.

Interaction partners of CikA from a yeast two-hybrid assay

Dr. Shannon R. Mackey used a yeast two-hybrid system to identify proteins that directly interact with CikA (Figure 7). Among five interacting CikA partners, she found two different conserved hypothetical proteins (Hypothetical protein 1 and Hypothetical protein 2) that interact with the PsR domain alone, which implicates these proteins uniquely as partners of the PsR (S. Mackey and S.S. Golden, manuscript in preparation).

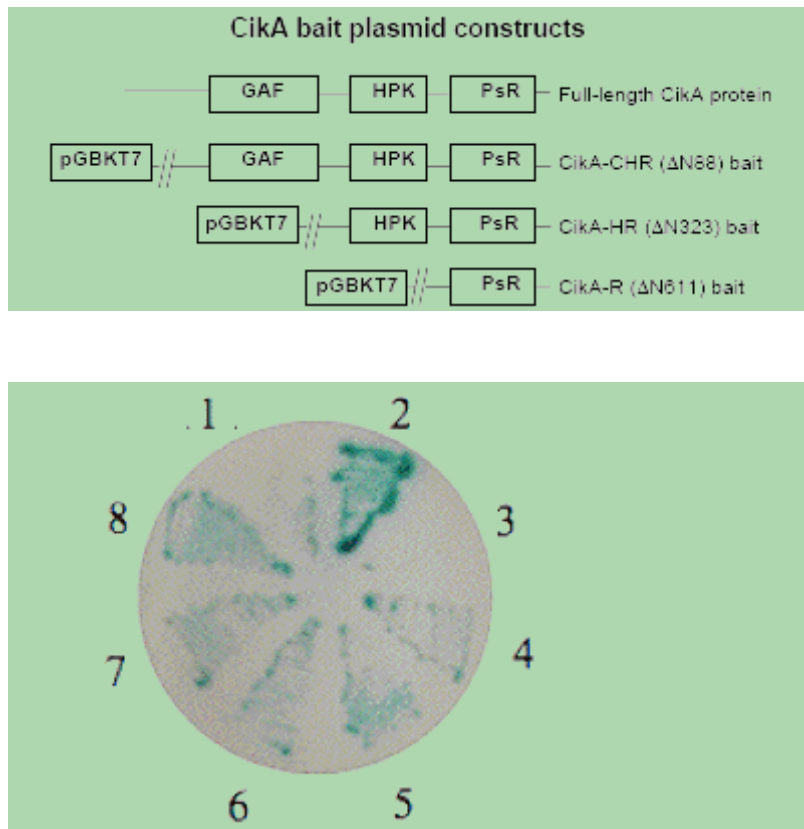


Fig. 7. Yeast two-hybrid assay. Three *cikA* bait plasmids were constructed to progressively delete one of the three predicted motifs: GAF, HPK, and PsR. The yeast two-hybrid screen was originally performed using Cika-HR. Potential partners of CikA in a Y2H screen. Growth and color indicate that the 'prey' clone encodes a protein that interacts with the 'bait.' 1- Cika-HR bait; 2- (+) control bait & prey (known interactors); 3- (-) control (known non-interactors); 4- NHT-1 prey; 5- Spk prey; 6- GAF-1 prey; 7 & 8- Hyp1 and Hyp 2; preys, which also interact with a PsR-only. (Figure and legend used, with permission, from S.R. Mackey and S.S. Golden).

Bioinformatics and genetics information regarding Hyp1

The Hyp1 sequence is not conserved among the completed cyanobacterial genomes. The N-terminal region of Hyp1 contains a predicted signal peptide sequence and a single transmembrane span; the C-terminus has a domain similar to the active domain of cytochrome c, which suggests it may bind a heme and be involved in sensing the redox state of the cell (Figure 10).

Overexpression of N-terminal 6XHis tagged Hyp1 fusion (N-6XHis-Hyp1) causes the waveform of the *psbAI* reporter to become more sinusoidal and less sharp than that of the WT control; also, the phase of peak expression occurs 8-10 h later than that of WT. Overexpression of 6XHis-Hyp1 in a *cikA* null causes a phase delay and damping of the rhythm. Hyp1 appears to be essential for viability, and complete disruption of Hyp1 would be lethal. In a yeast two-hybrid assay, the C-terminal half of Hyp1 interacts specifically with the PsR domain of CikA (S.R. Mackey and S.S. Golden, manuscript in preparation). Taking all of this information together, Hyp1 is a very important protein that is involved in the cyanobacterial circadian system. We predict that Hyp1 functions as a membrane-spanning protein that inserts its N-terminal transmembrane domain into the membrane at the poles and localizes CikA to the poles through interaction between the PsR domain of CikA and the C-terminal half of Hyp1.

Elucidating the role of Hyp1 will provide us with useful clues to the function and working mechanism of CikA, and therefore will help us to understand signal transduction in the clock system. In this graduate research, I used genetic, cell biological, and biochemical approaches to study the features and function of this newly identified clock component Hyp1. In this work I:

1. Tried to use an *in vitro* pull down assay to test interaction between Hyp1 and the PsR domain of CikA.
2. Constructed C-terminal half, N-terminal half, and full-length Hyp1 overexpression variants, fusing 6XHis to their C-terminal ends. The variants will be introduced into an *S. elongatus* PCC 7942 bioluminescent reporter strain by transformation and their overexpression phenotypes will be monitored by Dr. Shannon. R. Mackey using a luminometer. We expect the data will tell us the function of each part of Hyp1.

3. Constructed C-terminal half Hyp1-Yellow Fluorescence Protein (YFP) and full-length Hyp1-YFP overexpression fusions to investigate the sub-cellular localization of Hyp1. I got some preliminary data of localization, and Guogang Dong and Dr. Shannon R. Mackey in our lab will continue the analysis. Clarifying the subcellular localization of Hyp1 will not only test our hypothesis on Hyp1 function but also assist us in examining the inferred CikA function model.

In addition to that, I have conducted a project regarding KaiB. The crystal structure of KaiB strongly suggested that the functional complex is a tetramer formed as dimer of dimers, each composed of two non-equivalent subunits. The overall shape of this tetramer complex is an elongated hexagonal plate, and the tetramer has a positively charged cleft flanked by two negatively charged ridges made of terminal chains, suggesting that this cleft is a functional site (an active site or an interacting site) (25). If we could find some potential interactors of KaiB, it will help us to understand the function of KaiB. For this purpose, I used a yeast two-hybrid approach to look for interacting partners of KaiB. Unfortunately, I didn't find any positive interactors in my research. However, I will give a summary of the yeast two-hybrid assay for future reference by the Golden laboratory.

CHAPTER II

USING A PULL-DOWN ASSAY TO CONFIRM THE INTERACTION BETWEEN PsR AND HYP1 *IN VITRO*

Introduction

Hyp1 interacts with the PsR domain of CikA in a yeast two-hybrid *in vivo* assay. In this study, I tried to use a biochemical approach to confirm this interaction. The pull-down assay is an *in vitro* approach used to test protein-protein interactions. In a pull-down assay, a tagged bait protein is captured on an immobilized affinity ligand specific for the tag, thereby generating a secondary affinity support for other proteins that interact with the bait protein (38) (Figure 8).

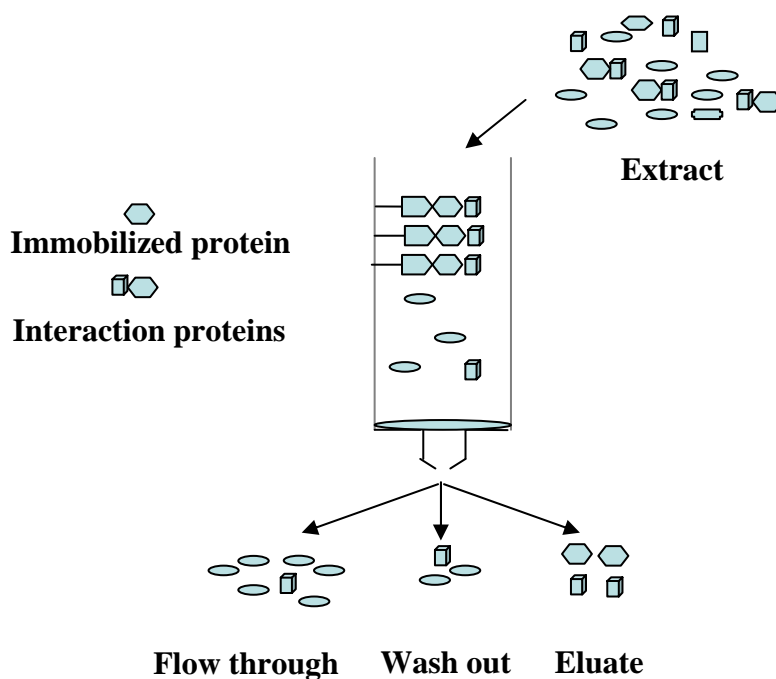


Fig. 8. Protein affinity chromatography. Extract proteins are passed over a column containing immobilized bait protein. Proteins that do not bind to the bait flow through the column, and ligand proteins that bind are retained. Strongly retained proteins will be eluted out after a solvent change to disrupt the interaction.

Bait proteins for pull-down assays can be generated by expressing recombinant fusion-tagged proteins. The bait can be incubated with a variety of other protein sources that contain putative prey proteins. When extract proteins are passed over a column containing immobilized affinity ligand, the bait protein will be captured and it will pull down a protein-binding partner (the prey). Proteins that do not bind to bait will either flow through the column or be washed out by the gentle washing buffer, and prey proteins that bind are retained. Identification of bait-prey interactions requires that the complex is eluted by the harsh elution buffer from the affinity support and analyzed by standard protein detection methods.

In all pull-down assays, carefully designed control experiments are required for generating biologically significant results. A negative control consisting of a non-treated affinity support (minus bait protein sample, plus prey protein sample) helps to identify and eliminate false positives caused by nonspecific binding of proteins to the affinity support. The immobilized bait control (plus bait protein sample, minus prey protein sample) helps to identify and eliminate false positives caused by nonspecific binding of proteins to the tag of the bait protein. The immobilized bait control also serves as a positive control to verify that the affinity support is functional for capturing the tagged bait protein.

In our experimental design, I tagged Hyp1 with 6 histidine residues (6XHis) as a bait protein. The Qiagen Ni-NTA matrix was used to capture the His-tagged Hyp1 on the column. The protein lysate containing overexpressed PsR was incubated with the lysate containing His-Hyp1 as well as Ni-NTA agarose. The negative control (without His-Hyp1) and immobilized bait control (without PsR) are described in the Materials and Methods section.

Materials and methods

Bacterial strains and culture conditions

Escherichia coli strain DH10B was used for routine cloning, and the *E. coli* BL21 (DE3) pLysE host strain (Novagen) was used for protein overexpression. The *E. coli* strains

hosting plasmids were cultivated in liquid or on agar-solidified LB medium in the presence of appropriate antibiotics (39).

DNA manipulations and construction of overexpression plasmids

QIAprep miniprep kit (QIAGEN) was used to prepare plasmid DNA for sequencing and subcloning. Sequencing was performed with BigDye terminator mix (Applied Biosystems) according to the manufacturer's instructions. Sequencing reactions were carried out by the Gene Technologies Laboratory (Texas A&M University). Other basic DNA manipulations were performed by standard procedures (39) using enzymes from New England Biolabs (Beverly, MA). PCR amplification was performed with home-made Pfu DNA polymerase. DNA fragments for subcloning were recovered from agarose gels using Promega Wizard SV Gel and PCR Clean-up System.

pET32a (Novagen), which encodes a thioredoxin (TRX) tag and cleavable 6XHis tag, was used as the overexpression vector for truncated Hypothetical 1 (Thyp1) cloning. The open reading frame that encodes Thyp1 (288-672 aa) was amplified by PCR with *S. elongatus* PCC 7942 genome as template. Two oligonucleotides, Truncated Hyp1 forward primer (5'-CGCGGATCCGATAACGTGGAGGGTTTGCATC-3') and Hyp1 Reverse primer (5'-CCCAAGCTTTTAAGTCTTCGGCGCTTTGTAG-3') incorporate a *Bam*HI site and *Hind*III site at its 5' termini, respectively. The resulting 1.1 kb PCR fragment was double digested with *Bam*HI and *Hind*III, and the digested fragment was purified and cloned into pET32a to generate the pET32a-His-Thyp1 construct. Once the DNA sequence was confirmed, the clone was introduced into *E. coli* BL21 (DE3) pLysE to allow expression of TRX-HIS-Thyp1 fusion under the control of the T7 *lac* promoter.

Quick-change mutagenesis of the PsR construct

The gene fragment that encodes residues 622-754 of CikA was cloned into the pET15b vector (Novagen) by Dr. Tiyu Gao in Dr. Andy LiWang's lab. The construct includes a T7 promoter and DNA sequence that encodes an N-terminal 6XHis tag. I used the plasmid pET15b-His-PsR to generate a non-His tagged mutant pET15b-PsR with the QuickChange Site-Directed Mutagenesis Kit (Stratagene). The gene-specific oligonucleotides, PsR QC For (5'-ACCATGGGCAGCAGCAGCAGCGGCCTGGTG-

3') and PsR QC Rev (5'-CACCAGGCCGCTGCTGCTGCTGCCCCATGGT-3') were used for Quickchange amplification reaction according to manufacturer's instructions.

Protein overexpression and analysis

Growth of cultures was performed as described in the Qiaexpressionist handbook (Qiagen). The only modification was using French press as a substitute for sonication.

E. coli BL21(DE3) pLysE containing pET32a-His-Thyp1 or pET15b-PsR was used to overexpress TRX-HIS-Thyp1 or PsR, respectively. A 5 ml overnight culture was used to inoculate 100 ml LB with appropriate antibiotics, and the culture was grown at 37°C shaking at 250 rpm until an OD600 of 0.6 was reached. A 1 ml sample was taken immediately before induction as the non-induced control. The cells were spun down and resuspended in 100 µl of 1X SDS gel-loading buffer (50 mM Tris•HCl, pH 6.8; 10% glycerol; 2% SDS; 4% mercaptoethanol, added just before use; 0.1% bromophenol blue). The culture was then induced with IPTG (1 mM, final concentration) for approximately 4 h. A second 1 ml sample was collected as an induced control and treated the same way as the first 1 ml sample. The rest of the cells were harvested by centrifugation at 4000 x g for 20 min and frozen at -80°C until ready for use.

To prepare the soluble lysate, the frozen *E. coli* pellets were thawed on ice for 15 minutes before being resuspended in 3 ml lysis buffer per gram pellet (50 mM NaH₂PO₄; 300 mM NaCl; 10 mM imidazole). Once the cells were resuspended, lysozyme was added (1 mg/ml, final concentration) and the cells were incubated on ice for 30 minutes. The cells were broken open by French press. The samples were applied to the French pressure cell and the cell was brought under the desired pressure (7000 to 10,000 psi). While maintaining the pressure, I adjusted the outlet flow rate to about one drop every second. Passage of the cell suspension was repeated twice and the cell lysate was collected in a flask kept on ice. After the cells were broken open, RNase (10 µg/ml, final concentration) and DNase (5 µg/ml, final concentration) were added and the cells were incubated on ice for 15 minutes. The cells were then collected by centrifugation at 10,000 x g for 30 minutes at 4°C to separate soluble and insoluble protein. I saved the supernatant fraction (soluble protein will be in the supernatant) at -80°C until ready for use.

To prepare the soluble protein for SDS-PAGE analysis, a 100 μ l cell lysate sample was mixed with equal volume of 2X SDS PAGE sample buffer (0.09 M Tris-HCl, pH 6.8; 20% glycerol; 2% SDS; 0.02% bromophenol blue; 0.01 M DTT).

The cell lysate, the frozen non-induced cells and induced cells were boiled for 5 min before loading on a gel. The prepared protein samples were separated on 12% or 15% SDS PAGE according to standard procedures (39).

To determine the optimal conditions to induce soluble Thyp1 expression, different IPTG concentrations (0.1 mM, 0.5 mM, and 1.0 mM), different induction times (2 h, 4 h, 6 h, 8 h, and 16 h) and different induction temperatures (18°C, 25°C, and 37°C) were combined for induction. The optimized condition was applied for overexpression induction when the volume was scaled up to 1 L.

Pull-down assay

Pull down procedures are very similar to purification protocols described in the Qiaexpressionist handbook (Qiagen). However, pull down should be performed using buffers with lower salt concentrations than used in buffers in simple purification of the target protein. Relatively low salt concentration does not disrupt protein interactions and allows co-purification of the target protein and its interactants.

Two main buffer systems were tried in this pull-down procedure. One buffer system was successfully used by Yong-Ick Kim in Dr. Andy Liwang's lab to co-purify His-tagged KaiA and untagged KaiC from *E. coli* (Interactant buffer: 150 mM NaCl, 20 mM Tris.Cl, 0.5 mM EDTA, 5 mM MgCl₂, pH 8.0; Washing buffer: 500 mM NaCl, 20 mM Tris. HCl, 500 mM imidazole; Elution buffer: 500mM NaCl, 20mM Tris. HCl, 500 mM imidazole). The other system was successfully used by Dr. Natasha B. Ivleva to co-purify proteins in cyanobacteria (Interactant buffer: 50 mM NaH₂PO₄, 5 mM NaCl. pH 7.8. Washing buffer: 50 mM NaH₂PO₄, 5 mM NaCl. 20 mM imidazole, pH 7.8; Elution buffer: 50 mM NaH₂PO₄, 5 mM NaCl. 250 mM imidazole, pH 7.8)

I induced 100 ml of *E. coli* BL21(DE3) pLysE containing pET32a-His-Thyp1 and pET15b-PsR separately as described above. I also induced 100 ml of *E. coli* BL21(DE3) pLysE containing pET32a vector only (no bait Thyp1) and pET15b vector only (no prey PsR) separately. The former one will be mixed with prey protein as a bait control. The

latter one will be mixed with bait protein as a negative control. To make it simple, pET32a plus PsR (control does not contain Thyp1), Thyp1 plus pET15b (control does not contain PsR) and His-Thyp1 plus PsR will be used in the following description.

I performed the following steps: (1) broke the cell with interactant buffers as described above; (2) mixed two lysates (three sets in total: His-Thyp1 plus PsR, His-Thyp1 plus pET15b, pET32a plus PsR) in a test tube with 1 ml Ni-NTA agarose, and mixed gently for 30min at room temperature; (3) loaded the mixture on an empty column and allowed them to flow through the column; (4) washed the column 5-12 times with 1ml washing buffer each time; (5) eluted the column 4 times with 1ml elution buffer each time; (6) collected each 1 ml fraction of flow-through, wash, and elution steps in a separate tubes (7) froze samples at -20°C until ready for use; (8) proceeded with analysis of wash and elution fractions with SDS-PAGE or immunoblot analysis.

Immunoblot analysis

Immunoblot analysis was performed as described earlier (10). Proteins were transferred to nitrocellulose membranes using electrophoretic transfer with a Trans-blot 321 SD semi-dry transfer cell according to the manufacturer's recommendations (Bio-Rad). His-tagged Thyp1 was detected using a Penta-His antibody (Qiagen) according to the manufacturer's directions and detected using a peroxidase-conjugated goat anti-mouse IgG (Jackson ImmunoResearch). For PsR detection, blots were blocked with TBS buffer supplemented with 0.1% Tween 20 and 1% milk, and then incubated for 2 h at room temperature with primary antiserum raised against CikA (1:7000 dilution). Antiserum was diluted in TBS buffer supplemented with 0.1% Tween 20 and 1% milk. After washing, blots were incubated for 1 h at room temperature with a 1:7000 dilution of peroxidase-conjugated goat anti-rabbit IgG (Calbiochem). The signal was visualized with the SuperSignal West Pico chemiluminescent substrate detection system (Pierce) and exposed to X-ray film.

Results

Bioinformatics analysis of Hyp1

The amino acid sequence of Hyp1 was compared to the sequences of other characterized cyanobacteria by standard search algorithms (BLASTP) (40). The results showed that it has 64% amino acid identity with the hypothetical protein sll1359 of *Synechocystis* sp. PCC 6803, but we cannot find a significant homolog in other completed cyanobacteria such as *Anabaena* sp. PCC 7120, *Thermosynechococcus elongatus* BP-1 or *Gloeobacter violaceus* PCC 7421, which indicates that Hyp1 is not universally conserved among cyanobacteria.

The N-terminal region of Hyp1 was predicted to contain a signal peptide sequence. Several programs (SMART, DAS, and TMpred) showed that a predicted single transmembrane domain is located at the region close to N-terminus. The C-terminus has a domain similar to the active domain of cytochrome c, which suggests it may bind a heme and be involved in sensing the redox state of the cell (Figure 9).

From the bioinformatics analysis, we predicted that Hyp1 is a membrane protein that may be involved in sensing the redox state of the cell.

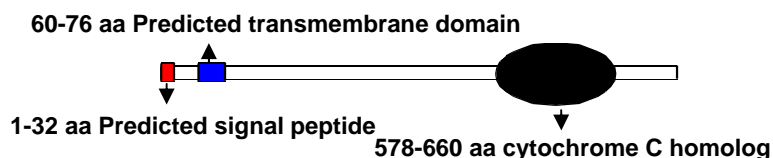


Fig. 9. Bioinformatics analysis of Hyp1. The N-terminal region of Hyp1 was predicted to contain a signal peptide sequence (red box) (signal peptide probability 0.986 with cleavage site probability 0.677 at residue 32 using the Signal P 2.0 HMM program (1). Several programs (SMART, DAS, and TMpred) showed that a predicted single transmembrane domain (Blue box) is located at the region close to N-terminus. The C-terminus has a domain similar to the active domain of cytochrome c (black circle).

Optimizing truncated Hyp1 protein overexpression

The DNA fragment that encodes full-length Hyp1 (1-672 aa) was cloned into overexpression vectors pQE30 and pET32a, sequentially, but no protein overexpression could be detected from either plasmid with SDS-PAGE analysis. I truncated the N-terminal region of Hyp1 and removed the predicted signal peptide and transmembrane domain. Only the C-terminal half of Hyp1 (288-672 aa), which specifically interacts with PsR in a yeast two-hybrid assay, was encoded by the sequence cloned into pET32a to yield a thioredoxin-fused protein (pET32a-His-Thyp1). Sequence analysis showed that it contained the proper open reading frame (ORF) in-frame and under T7 promoter control in the vector. With 1 mM IPTG induction, the expected 60.9-kD TRX-HIS-Thyp1 fusion protein was detected by SDS-PAGE (Figure 10), and a dramatic increase of expression after 1mM IPTG induction can be seen clearly (comparing lanes 3 and lane 4). This strain was preserved for later use in the pull-down assay.

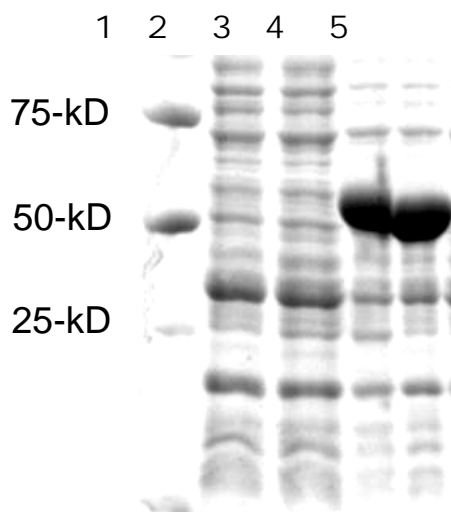


Fig. 10. Overexpression of TRX-HIS-Thyp1 (expected size: 60.9 kD). After 1mM IPTG induction for 4 hours at 37°C, whole cell proteins were loaded onto SDS-PAGE to examine TRX-HIS-Thyp1 overexpression. 1. Protein Marker. 2. Negative control (pET32a in *E.coli* BL21(DE3) pLysE). 3. pET32a-Thyp1(non-induced control). 4. induced pET32a-Thyp1. 5. Positive control (pET32a-CikA CBD in *E.coli* BL21 (DE3) pLysE).

Once the protein can be overexpressed in the presence of inducer, the problem of solubility needs to be addressed. With 1mM IPTG induction for 4 hours at 37°C, it is hard to detect soluble TRX-HIS-Thyp1 signal in the soluble protein lysate with Coomassie Brilliant Blue staining (data not shown), which indicates that most of the TRX-HIS-Thyp1 is insoluble. Many factors affect protein solubility including growth temperature, concentration of inducer, and length of induction. To optimize the conditions for the soluble protein expression, many different trials were executed (see Materials and Methods). The soluble fraction of the lysate from each trial was applied to SDS-PAGE to examine the soluble protein level. Induction under 37°C or 25°C did not produce much soluble TRX-HIS-Thyp1 (data not shown). Eventually, I reduced the induction temperature to 18°C and tested different IPTG concentrations and induction durations. The result in Figure 11 shows that the cells induced with 0.1mM IPTG for 12 hours at 18°C (lane 3) will give maximum soluble protein. This optimized induction condition was used for large-scale induction later.

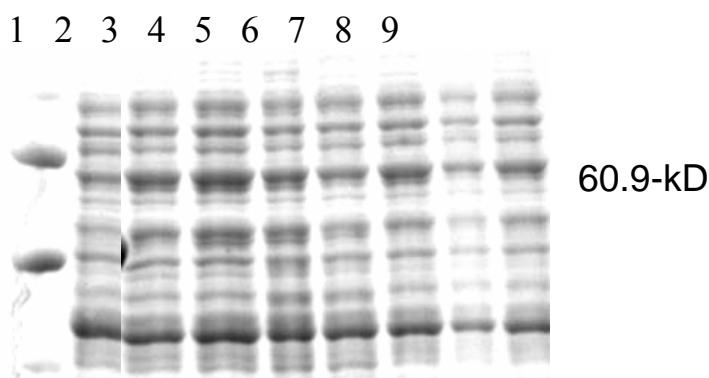


Fig. 11. Thyp1 protein expression was optimized to get maximum soluble protein. To optimize the conditions for soluble protein expression, many different trials were executed. Different IPTG concentrations and different induction times were combined to induce Thyp1 overexpression. The same amount of induced cells was harvested and broken with a French press and the supernatant fraction was mixed with an equal volume of 2XSDS PAGE sample buffer for loading. Lysate from *E. coli* BL21 (DE3) pLysE containing pET32a-Thyp1 which was (1) uninduced and induced with (2) 0.1 mM IPTG for 6 h at 18°C; (3) 0.1 mM IPTG for 12 h at 18°C; (4) 0.1 mM IPTG for 16 h at 18°C; (5) 0.5 mM IPTG for 2.5 h at 18°C; (6) 0.5 mM IPTG for 6 h at 18°C; (7) 1 mM IPTG for 2.5 h at 18°C; or (8) 0.1 mM IPTG for 4 h at 18°C.

Overexpression of PsR without a 6XHis tag

We obtained a His-tagged PsR overexpression construct (pET15b-His-PsR) from Dr. Andy Liwang's lab at Texas A&M University. Because one of the partners to be used in the pull-down assay, Thyp1, already has a 6XHis tag, we wanted to eliminate the 6XHis from PsR. Initially, I used thrombin, a specific protease, to cut the 6XHis from 6XHis-PsR fusion protein. However, this approach caused some trouble in pull-down assays because we cannot ensure that the 6XHis is completely removed from the fusion protein; there is only a slight size difference between His-PsR and PsR and they separate poorly by SDS-PAGE. Any remaining His-PsR would be retained on a Ni-NTA column and show up in the eluate independent of the presence of Thyp1. Eventually, I used the Quickchange approach to remove the 6XHis genetically. The sequencing analysis confirmed that the codons for the 6XHis were gone. Overexpression of PsR without a 6XHis and its solubility were examined on SDS-PAGE. A large amount of soluble PsR was detected in the lysate (Figure 12). Comparing lanes 6 and lane 7, we can see a size different between the His-PsR fusion and PsR alone. This is additional evidence that the 6XHis is indeed removed.

Co-purification of PsR and Thyp1

In the pull-down assay, TRX-HIS-Thyp1 was used as a bait, and PsR was considered as a prey. In order to test the hypothesis of Thyp1-PsR interaction, recombinant 60.9-kD TRX-HIS-THyp1 and 14.6-kD non-His-tagged PsR were expressed in *E. coli*. Cell lysates (lysed in Yong-Ick Kim's reaction buffer, see Methods and Materials) of induced bacteria were mixed together with Ni-NTA agarose and incubated for 30 min. The mixture was applied to an empty column. To eliminate the false positives and false negatives, cell lysates that do not contain the prey protein or bait protein were used as controls (see Materials and Methods).

In the pET32a plus PsR control, we predict that PsR cannot be retained on the column since it has no 6XHis. It will be detected in the flow through or wash fractions, but not in the eluate. However, in the His-Thyp1 plus PsR, if PsR interacts with Thyp1, PsR will be captured by His-Thyp1 on the column, and can be detected in the eluate.

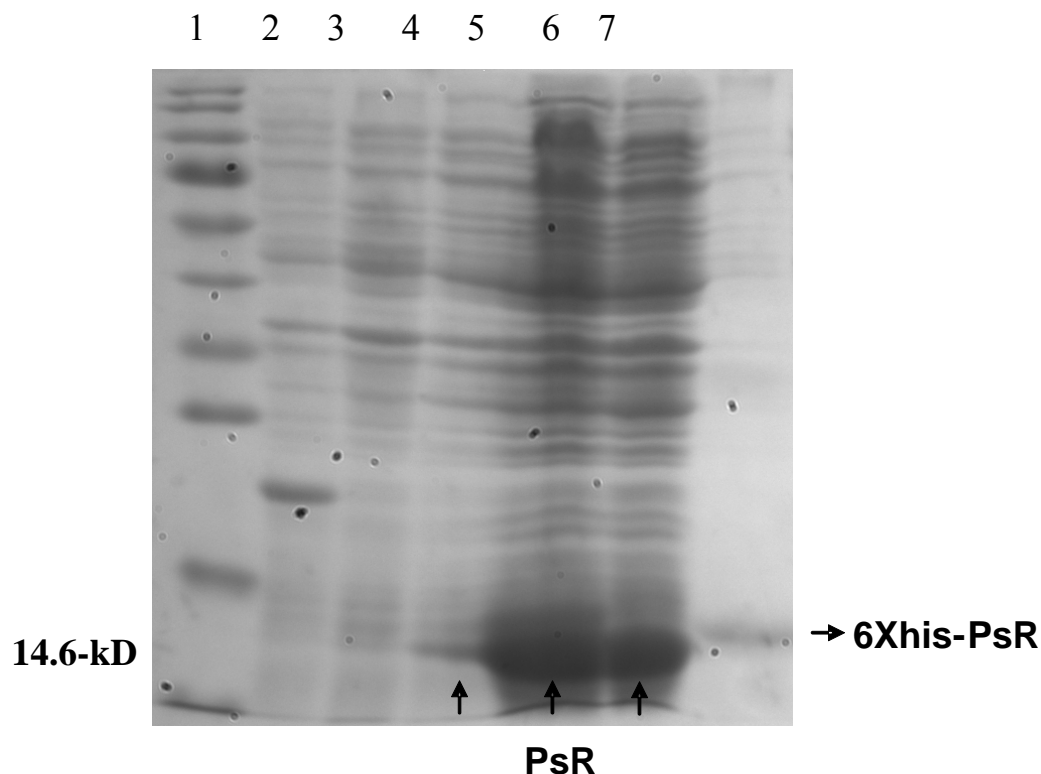


Fig. 12. The overexpression of PsR after genetically removing the 6XHis. The induced cells were treated as described in Materials and Methods. The whole cell protein and soluble purified protein were loaded on an SDS-PAGE gel. 1. Protein Marker. 2. Negative control (pET32a in *E.coli* BL21 (DE3) pLysE). 3. Non-induced control: pET15b-PsR in *E. coli* BL21 (DE3) pLysE. 4. Induced control: pET15b-PsR in *E. coli* BL21 (DE3) pLysE. 5. Soluble PsR, 10 ul for loading. 6. Soluble PsR, 5 ul for loading. 7. Purified His-PsR (purified on Ni-column).

The proteins bound to the matrix were recovered after successive washes. SDS-PAGE followed by Western blot analysis (Figure 13) showed that PsR was detected in the eluate not only in the Thyp1 plus PsR sample, which is consistent with our hypothesis (lane 6), but also in the PsR plus pET32a (lane 5), in which Thyp1 is absent. PsR showing up in eluate independent of Thyp1 suggests that PsR itself will bind to the Ni-NTA column non-specifically. I increased the washing times and washing stringency to repeat the pull-down experiments several more times. However, when washing stringency is too high, PsR cannot be detected in the eluate from either the PsR plus pET32a or PsR plus

Thyp1 samples. When washing stringency was too low, PsR showed up in both situations (Data not shown).

The working solutions for washing and binding are physiological in pH and ionic strength, providing a starting point from which specific buffer conditions for each unique interacting pair can be optimized. Because Yong-Ick Kim's buffer system did not work for these assays, I used Dr. Natasha B. Ivleva's buffer system to do pull-downs again. I gradually increased the washing stringency by increasing the imidazole concentration and adjusting washing times, and all the fractions of washing and eluting steps were collected. Here, I show three representative experimental results:

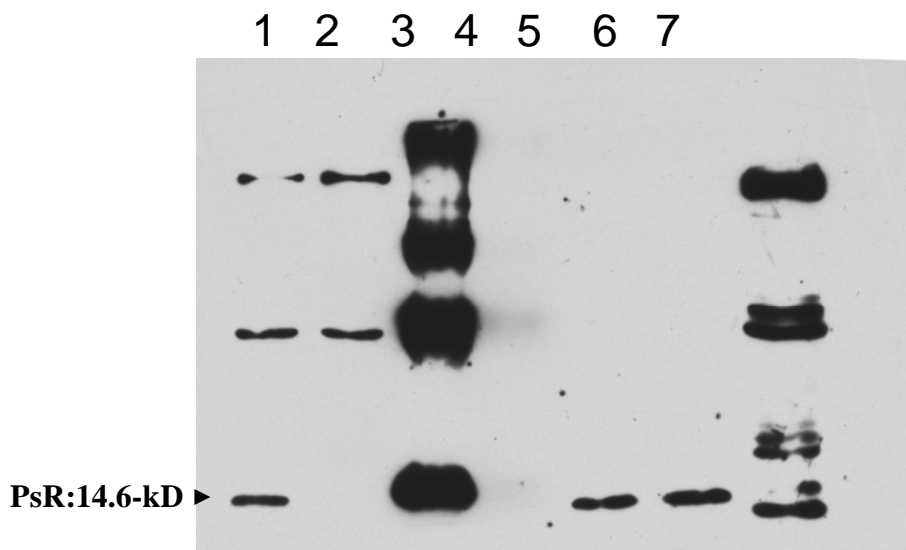


Fig. 13. PsR was detected in the eluate independent of presence of His-Thyp1. Relevant lysates were mixed in a test tube and loaded on columns (See Materials and Methods). The eluting fractions were examined with SDS-PAGE and immunoblot analysis using antiserum raised against CikA. 1. Positive control: cell lysate of *E.coli* BL21 (DE3) pLysE containing pET15b-PsR. PsR is present but Thyp1 is absent. 2. Negative control: cell lysate of *E. coli* BL21 (DE3) pLysE containing pET32a. Both PsR and Thyp1 are absent. 3. Positive control: cell lysate of *E. coli* BL21 (DE3) pLysE containing pET15b-His-PsR. His-PsR is present, but Thyp1 is absent. 4. Eluate of pET32a plus THyp1. 5. Eluate of pET32a plus PsR. 6. Eluate of THyp1 plus PsR. 7. Negative control: cell lysate of *E.coli* BL21 (DE3) pLysE containing pET32a-TXR-HIS-Thyp1. His-Thyp1 is present but PsR is absent.

Under low stringency washing (washing buffer: 50 mM NaH_2PO_4 , 5 mM NaCl, 10 mM imidazole; wash 10 times, 1 ml each time), PsR was detected in both wash and elution fractions. There was no difference between the amounts of PsR eluted in the Thyp1 plus PsR and pET32a plus PsR samples (Figure 14A).

Under high stringency washing (washing buffer: 50 mM NaH_2PO_4 , 5 mM NaCl, 20 mM imidazole; wash 10 times, 1ml each time), no PsR was detected in the tenth wash fraction (W10) and elution fractions in both the Thyp1 plus PsR and pET32a plus PsR samples (Figure 14B).

Under medium stringency washing (washing buffer: 50 mM NaH_2PO_4 , 5 mM NaCl, 15 mM imidazole, wash 6 times, 1ml each), PsR was detected in both wash and elution fractions. Although we can see some differences between the pET32a plus PsR and Thyp1 plus PsR samples, the quantitative differences between them are not sufficient for us to draw the conclusion that there is an interaction between Thyp1 and PsR (Figure 14C).

These results showed that it was very hard to find an optimal washing stringency that allows us to exclude the non-specifically bound PsR.

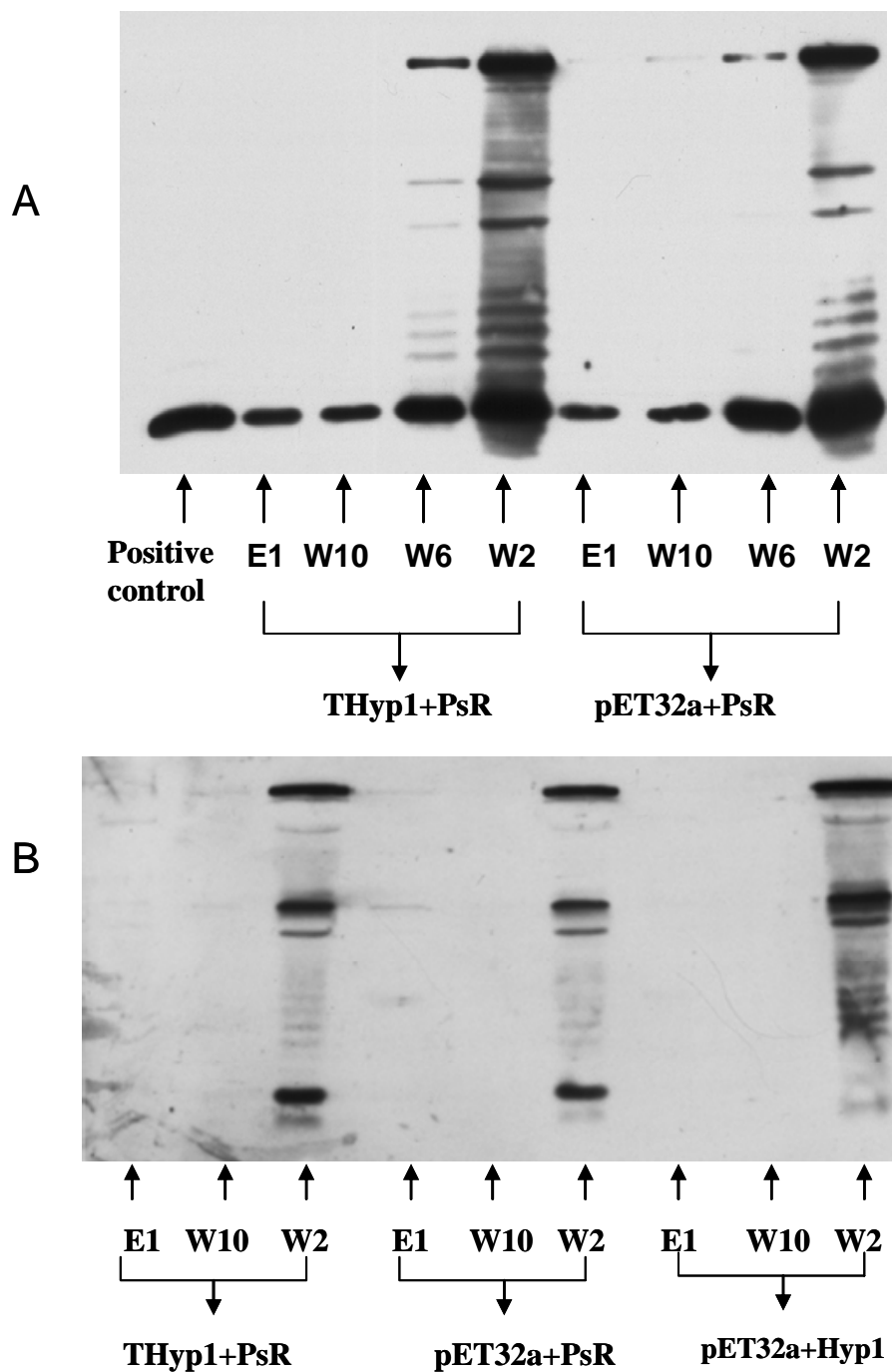


Fig. 14. Pull-down assays with different washing stringencies. Changing washing stringencies to exclude the non-specifically bound PsR. Fractions of flowing through, washing, and eluting steps are analyzed with SDS-PAGE and immunoblot analysis. A. Low washing stringency. B. High washing stringency. C. Medium washing stringency. Positive control: Cell lysate of *E. coli* BL21 (DE3) pLysE containing pET15b-PsR. Negative control: Cell lysate of *E. coli* BL21 (DE3) pLysE containing pET32a. E1, E2, E3: Eluting fraction 1, 2, 3. W2, W4, W6, W10: washing fraction 2, 4, 6, 10.

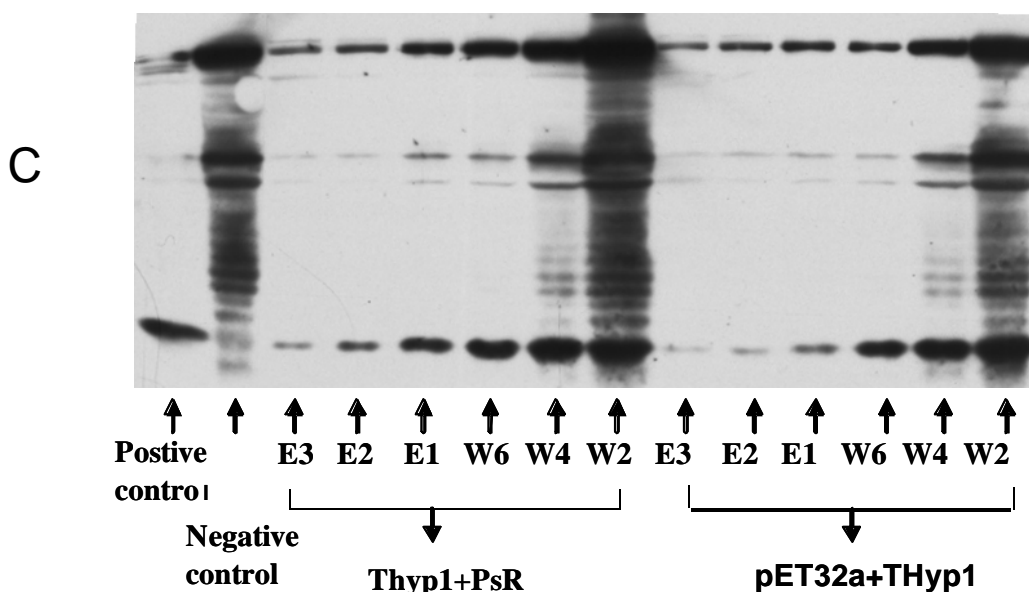


Fig. 14. Continued

Discussion

The yeast two-hybrid assay demonstrated the interaction between CikA and Hyp1 *in vivo*. In this study, I tried confirming this interaction through biochemical approach *in vitro*. We found that Hyp1 has very low solubility, which is consistent with our hypothesis that Hyp1 is a membrane protein. When I removed the N-terminal half of Hyp1 and optimized the induction condition, I obtained soluble truncated Hyp1. These results are consistent with the hypothesis that Hyp1 is the membrane protein that is responsible for localizing CikA at the membrane near the pole of the cell.

I was surprised to find that non-His tagged PsR can bind non-specifically to a Ni-NTA column. I tried to optimize the washing stringency to allow selective elution of prey protein while the bait remains immobilized. However, using a step-wise gradient of increasing imidazole concentration and changing buffer system did not help.

Although our pull-down assay was inconclusive, the results do not necessarily mean that PsR cannot interact with Thyp1. Since PsR will nonspecifically bind to Ni-NTA column for unknown reasons, in future studies we can use other affinity chromatography methodologies instead. For an instance, we can subclone the Hyp1 gene

to an appropriate vector with a glutathione S-transferase tag (GST-tag), and then use the appropriate affinity chromatography for pull-down assay. An alternative way to do *in vitro* protein-protein interaction assay is by using co-immunoprecipitation. I have purified Thyp1 to prepare Hyp1 antibody, and PsR can be detected with anti-CikA. We can add antibody specific to Hyp1 to a cell lysate containing both Hyp1/Thyp1 and PsR. Then the antibody-protein complex is precipitated using protein-G sepharose which binds most antibodies. If PsRs bind to Hyp1, they will also be pelleted.

Discovery and confirmation of protein-protein interactions using the pull-down technique depend heavily on the nature of the interaction under study. Interactions can be stable or transient and this characteristic determines the conditions for optimizing binding between bait and prey proteins. Stable protein-protein interactions are easier to isolate by physical methods like pull-down assays because the protein complex does not disassemble over time. Unstable, transient interactions are the most challenging protein-protein interactions to isolate. These interactions are more difficult to identify using physical methods like pull-down assays because the complex may dissociate during the assay. Some transient interactions occur during transport or as part of enzymatic processes; they often require cofactors and energy via nucleotide triphosphate hydrolysis. Incorporating these factors during an assay will be essential for success if such molecules are usually involved. Currently, we have limited information about the features and function of Hyp1: we neither know the nature of interaction between Hyp1 and PsR, nor do we know if any cofactors are required for their interaction. For this reason, I propose that it will be safer to use a cyanobacterial system instead of *E. coli* to do the *in vitro* pull-down assay because, in cyanobacteria, most of the potential interaction cofactors will be kept.

CHAPTER III

CONSTRUCTION OF THREE HIS-TAGGED HYP1 VARIANTS AND TWO HYP1-YELLOW FLUORESCENT PROTEIN FUSIONS FOR GENETIC AND SUBCELLULAR LOCALIZATION ANALYSES OF HYP1

Introduction

In the inferred structure and function model of CikA, CikA is thought to be located at the poles through interaction between PsR and some protein or protein complex that is associated with the cell membrane and also localized at the poles (33). Five potential interactors have been identified through a yeast two-hybrid screen, of which two hypothetical proteins, Hyp1 and Hyp2, specifically recognize a PsR domain bait (S.R. Mackey and S.S. Golden, manuscript in preparation). No recognizable motifs can be found in Hyp2 through bioinformatics analysis. However, the bioinformatics analysis did show that there is a predicted signal peptide and a transmembrane span at the N-terminal of Hyp1 and a cytochrome C homolog (heme-binding) domain at the C-terminal of Hyp1. What is the function of the N-terminal and C-terminal parts of Hyp1? Is Hyp1 a membrane protein which is located at the poles? Does Hyp1 determine the proper localization of CikA? Does the interaction between Hyp1 and PsR fulfill the functions proposed in the CikA function model? To infer the function of each part of Hyp1, I constructed three Hyp1-encoding alleles which will overexpress the C-terminal half of Hyp1, the N-terminal half of Hyp1, or full-length Hyp1, separately in cyanobacterial reporter strains. Circadian phenotypes will be monitored by the oscillation in bioluminescence on a TopCount luminometer (Perkin-Elmer) for all the constructs by Dr. Shannon R. Mackey.

The validity of using YFP as a fusion tag to localize circadian clock proteins in *S. elongatus* has been confirmed (unpublished data in our lab, G. Dong and S.S. Golden). To test the hypothesis that Hyp1 is a transmembrane-spanning protein that docks CikA at the poles through interaction between PsR and the C-terminus of Hyp1, I fused YFP in

frame with full-length Hyp1 and with the C-terminal half of Hyp1 at the C-terminal end of each variant. In this chapter, I will describe how these overexpression plasmids were constructed and what we can do with these constructs in the future.

Materials and methods

Bacterial strains and plasmids

The plasmids used in this study are listed in Table 1. All plasmids were maintained in *E. coli* DH10B cells. The media and growth conditions are described in Methods and Materials in Chapter II. Wild-type and mutant strains of the cyanobacterium *S. elongatus* PCC 7942 were grown at 30°C in BG-11M medium as 100 ml cultures shaken constantly at 250 r.p.m. under 300 $\mu\text{E m}^{-2} \text{s}^{-1}$ of fluorescent light as described previously (32). AMC1084 and AMC1085, which were transformed by pAM3760 and pAM3766 respectively, were grown in the presence of spectinomycin (Sp, 20 mg/ml).

Constructing three C-terminal His-tagged Hyp1 overexpression variants

DNA manipulations and sequencing are described in Materials and Methods in Chapter I. Cyanobacterial transformations were performed as described earlier (7).

pAM2991 was used as an overexpression vector for cloning His-tagged Hyp1 overexpression variants. The segments of the ORF that encode the C-terminal half of Hyp1 (288-672 aa), the N-terminal half of Hyp1 (1- 285 aa), and full-length Hyp1(1-672 aa) were amplified by PCR with *S. elongatus* PCC 7942 genomic DNA as template. The 6XHis tag encoding sequence and appropriate restriction site were incorporated in the PCR primers (listed in Table 2).

Constructing three C-terminal His-tagged Hyp1 overexpression variants

DNA manipulations and sequencing are described in Materials and Methods in Chapter I. Cyanobacterial transformations were performed as described earlier.

Table 1: Plasmids used in this study

| Plasmids | Characteristics | Antibiotic resistance | Source or Reference |
|-----------------|--|------------------------------|--------------------------------|
| pAM2991 | NS1 cloning vector; IPTG-inducible promoter | Sp/Sm | Ivleva, et al. 2005 |
| pAM3699 | NS1 cloning vector; IPTG-inducible promoter | Sp/Sm | Guogang Dong Lab collection |
| pAM3760 | Truncated <i>hyp1-yfp</i> fusion; overexpression construct in pAM3699 | Sp/Sm | This study |
| pAM3762 | N-terminal half of <i>hyp1</i> ; overexpression construct in pAM2991 | Sp/Sm | This study |
| pAM3763 | C-terminal half of <i>hyp1</i> ; overexpression construct in pAM2991 | Sp/Sm | This study |
| pAM3765 | Full-length <i>hyp1</i> overexpression construct in pAM2991 | Sp/Sm | This study |
| pAM3766 | Full-length <i>hyp1-yfp</i> fusion; overexpression construct in pAM3699 | Sp/Sm | This study |

pAM2991 was used as an overexpression vector for cloning His-tagged Hyp1 overexpression variants. The segments of the ORF that encode the C-terminal half of Hyp1 (288-672 aa), the N-terminal half of Hyp1 (1- 285 aa), and full-length Hyp1(1-672 aa) were amplified by PCR with *S. elongatus* PCC 7942 genomic DNA as template. The 6XHis tag encoding sequence and appropriate restriction site were incorporated in the PCR primers (listed in Table 2).

The resulting PCR fragments were digested with appropriate restriction enzymes, and the digested fragments were purified and cloned into pAM2991 to generate overexpression constructs.

Constructing full-length *hyp1-yfp* and truncated *hyp1-yfp* overexpression alleles

A three-way ligation strategy was used to tag full-length Hyp1 and truncated Hyp1 with YFP at the C-terminus. The portions of the ORF that encode full-length Hyp1 (1-672 aa) and truncated Hyp1 (288-672 aa) were amplified with primers that incorporate *NcoI* and *XhoI* restriction sites at upstream and downstream ends, respectively. The YFP ORF was amplified with primers that incorporate *XhoI* and *BamHI* restriction sites at upstream and downstream ends, respectively. All sequences of primer are listed in Table 2. The resulting PCR fragments were digested with appropriate restriction enzyme. The plasmid pAM3699 was digested with *NcoI* and *BamHI*. The proper digested fragments were purified and ligated together overnight. The resulting ligation product was introduced into *E.coli* DH10B. The fluorescent positive clones were picked up for sequencing.

Table 2: Primers used for cloning

| |
|---|
| 1. Full length Hyp1 Forward primer: |
| 5'-CCG <u>GAATTC</u> ATGTCAAAACAAAAATGGGCCG-3' |
| <i>EcoRI</i> |
| 2. Full length Hyp1 Reverse primer: |
| 5'-CGC <u>GGATCCTTAA</u> <u>ATGGTGATGGTGATGGTG</u> AGTCTTCGGCGCTTTTAG-3' |
| <i>BamHI</i> 6XHis |
| 3. N-Hyp1 Reverse: |
| 5'-CGC <u>GGATCCTTAA</u> <u>ATGGTGATGGTGATGGTG</u> TTCATAGGTACTCCAAATCCACTGG-3' |
| <i>BamHI</i> 6XHis |
| 4. C-Hyp1 Forward: |
| 5'-CCG <u>GAATTC</u> ATGGATAACGTGGAGGGTTTGCATCC-3' |
| <i>EcoRI</i> |
| 5. Gd043f (<i>xhoI</i> YFP forward) |
| 5'-GG <u>CTCGAG</u> TATGGTGAGCAAGGGCGAGGAG-3' |
| <i>XhoI</i> |
| 6. Gd043r (<i>BamHI</i> YFP reverse) |
| 5'-GACT <u>GGATCC</u> TTACTTGTACAGCTCGTCCATGC-3' |
| <i>BamHI</i> |
| 7. <i>NcoI</i> Hyp1 For: |
| 5'-GCCA <u>CCATGG</u> CAATGTCAAAACAAAAATGGGCC-3' |
| <i>NcoI</i> |
| 8. <i>XhoI</i> Hyp1 Rev: |
| 5'-GACT <u>CTCGAG</u> CCAGTCTTCGGCGCTTTGTAG-3' |
| <i>XhoI</i> |
| 9. <i>NcoI</i> Thyp1 For: |
| 5'-GCCA <u>CCATGG</u> CAGATAACGTGGAGGGTTTG-3' |
| <i>NcoI</i> |

Notes. I paired primers (primer X/primer Y) to amplify: (1/2) full-length *hyp1*; (1/3) N-terminal coding region of *hyp1*; (2/4) coding region of *hyp1*; (5/6) *yfp*; (7/8) full-length *hyp1*; (8/9) truncated *hyp1*.

Visualization of intracellular localization of YFP

To facilitate imaging, YFP fusions were induced by adding a final concentration of 1 mM IPTG to fresh cultures of reporter strains and fluorescence was observed about 10 h later. Samples were loaded directly on a slide treated with poly-L-lysine, covered with a cover slip, and sealed with nail polish. Differential Interference Contrast (DIC) images were acquired with a Zeiss Axioplan2 microscope and Hamamatsu C5810 3CCD camera. Exposure time was 1 s for fluorescence images and 0.33 s for DIC. All images were processed with the publicly-available program ImageJ (<http://rsb.info.nih.gov/ij>). A convolution filter with default settings was applied to the red autofluorescence image arising from chlorophyll to achieve apparent sectioning of the cell and reveal the peripheral localization of the photosynthetic apparatus.

Results and discussion

Constructing three C-terminal His-tagged Hyp1 overexpression variants

Gene disruption and gene overexpression are two commonly used strategies to get functional insights into a gene. Constructing a series of deletion and overexpression variants will help us to examine the function of each domain or bioinformatics unit of a gene. It is impossible to segregate null alleles of *hyp1* in *S. elongatus*, which suggests that Hyp1 is essential for viability (S.R. Mackey and S.S. Golden, manuscript in preparation). To investigate Hyp1's function, Dr. Shannon R. Mackey overexpressed full-length Hyp1 with a 6XHis-tag at its N-terminal. This increased amount of Hyp1 protein caused the waveform of the bioluminescence rhythm from a *psbAI*-luciferase reporter to become more sinusoidal and less sharp than that of the WT control; also, the phase of peak expression occurs 8-10 h later than that of WT. Overexpression of 6XHis-Hyp1 in a *cikA* null causes a phase delay as well as damping of the rhythm (S.R. Mackey and S.S. Golden, manuscript in preparation). A quantification of Hyp1 expression level is complicated by the fact that the N-terminal 6XHis tag apparently was cleaved off in *S. elongatus*, which is consistent with the bioinformatics prediction that there is a signal peptide at N-terminal (S.R. Mackey and S.S. Golden, manuscript in preparation). The N-terminal half of Hyp1 and C-terminal half of Hyp1 have different bioinformatics

characteristics. To study each part's function *in vivo*, I constructed three overexpression variants with a 6-His tag fused to the C-terminus. Each construct is schematically represented in Figure 15. All constructs were confirmed by sequencing to be in-frame and inserted under control of a *P_{trc}* promoter, which is inducible by isopropyl- β -D-thiogalactopyranoside (IPTG). The *P_{trc}* promoter exhibits low-level basal expression in the cyanobacterium (32). We could test both complementation in uninduced conditions and an overexpression phenotype in the presence of inducer, for all constructs.

All constructs will be introduced into wild-type *luxAB* reporter strain of *S. elongatus* and integrated at a neutral site (NS I) in the chromosome. The circadian phenotypes under induced and un-induced conditions will be monitored on a TopCount luminometer by Dr. Shannon R. Mackey. Hyp1 carries a motif characteristic of a cytochrome C in its C-terminal half. It may be involved in sensing the redox state of the cell. Evidence shows that both LdpA and CikA bind redox-active cofactors and sense light indirectly through sensing the redox state of the cell (29, 34). I propose that Hyp1 is a membrane protein with its N-terminal half inserted into the cell membrane that can pass redox information collected by a heme group in its C-terminal half to CikA through interaction between PsR and C-terminal half of Hyp1. In addition, this interaction will dock CikA to the cell pole for proper function.

I predict that overexpression of full-length Hyp1 with a His tag at its C-terminal in a wild-type reporter strain will result in the same phenotype as the N-6-His tagged Hyp1. It is very hard to predict what kind of phenotype we can observe when producing the C-terminal half or N-terminal half only in excess in a wild-type reporter strain because maybe C-terminal function will rely on the N-terminal domain for proper localization. If these two parts function independently, when the C-terminal half is overexpressed, most likely it will be in the cytoplasm, and will compete with wild-type Hyp1 for PsR binding. Therefore, CikA will not localize at the poles for proper functioning, and overexpression will result in a phenotype similar to that of a PsR null strain. Overexpressing only the N-terminal half of Hyp1 in the wild-type strain may have a dominant negative effect in the cell over the wild-type copy of Hyp1. Because N-terminal Hyp1 has a transmembrane domain, the extra N-terminal Hyp1 proteins are proposed to insert into the membrane and compete for the localization of wild-type Hyp1.

Therefore, fewer wild-type Hyp1 proteins will localize on the poles. Because the N-terminal Hyp1 at the poles have no PsR-interacting domain, most of the CikA will lose its proper localization and function.

If we do not observe any abnormal phenotypes of overexpression in wild-type reporter strains, we can disrupt the wild type copy of Hyp1 in the genome to observe the cell viability. If any part of Hyp1 (C-terminal or N-terminal half) will rescue the Hyp1-disruption phenotype, it will indicate that this part is essential for cell viability. The circadian phenotype of rescued strains will give us clues to the function of that part of the protein.

After we get more information about features and functions of each part, we can introduce these constructs into *cikA* or *kai* null strains to see their dependence on the other circadian proteins.

Construction and overexpression of full-length *hyp1-yfp* and truncated *hyp1-yfp* fusion alleles

The most direct test of our hypothesis of Hyp1 function is *in vivo* imaging of the Hyp1 protein. Green Fluorescent Protein (GFP) is one of the most popular reporters used for intracellular localization analysis, but, for some reason, it doesn't fluoresce in *S. elongatus*. Guogang Dong in our lab has successfully created some fluorescent reporters with other fluorescent proteins— ZsGreen (ZsG), a coral reef fluorescent protein, and a yellow variant of the green fluorescent protein (YFP)— to examine the localization of Kai proteins and CikA (33). In this study, I created a truncated Hyp1-YFP fusion reporter construct and a full-length Hyp1-YFP fusion reporter construct (Figure 15) to investigate the sub-cellular localization of Hyp1.

Both overexpression constructs were confirmed by sequencing, and the ORFs are inserted in-frame under control of the *P_{trc}* promoter. Both constructs were introduced into *S. elongatus* by transformation and integrated at a neutral site (NS1) in the chromosome (41). According to our hypothesis, Hyp1 is a transmembrane-spanning protein that docks CikA at the poles through interaction between PsR and the C-terminus of Hyp1. I predict that the full-length Hyp1 will localize at the cell poles and the C-

terminal half of Hyp1 will distribute in the cytoplasm because the transmembrane domain of Hyp1 is removed.

Unfortunately for us, the *P_{trc}* promoter exhibited basal expression in the *E. coli* DH10B in this study. The YFP tagged to Thyp1 glows well in *E. coli* DH10B even without IPTG induction, while it glows relatively weakly in the full-length Hyp1 construct. When they were introduced into wild-type *S. elongatus*, and after chromosomal segregation and clonal propagation, the IPTG-induced cyanobacterial cells were monitored under the fluorescence microscope for Hyp1 localization. The preliminary data showed that most of the cells did not fluoresce brightly after IPTG induction. Only a few cells glowed brightly enough to make it possible to track the localization of Hyp1, but the localization varied in different cells. Therefore, we cannot draw any conclusions from these data.

In the future, we can either add a longer linker between Hyp1 and YFP to improve the likelihood of preserving Hyp1 function or put the fusion under control of a different promoter to see if we can get more consistent overexpression in cyanobacteria. After we get the localization information for Hyp1, we can go ahead to see if this kind of localization is CikA dependent or not in a *cikA* null strain.

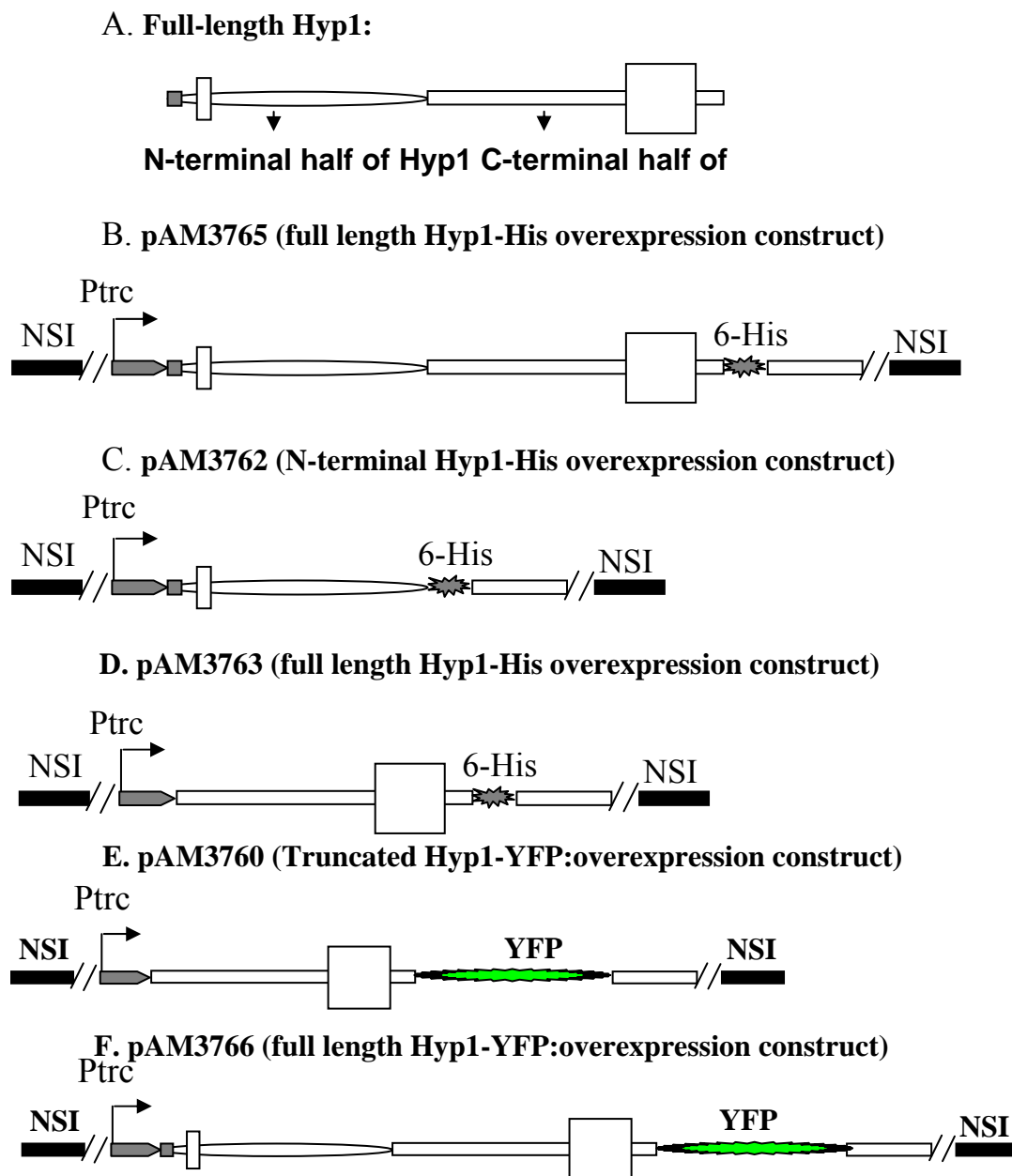


Fig. 15. Schematic representation of each construct. A. Schematic representation of full length Hyp1. B. Schematic representation of full length Hyp1 of pAM3765 (full length Hyp1-His overexpression construct). C. Schematic representation of pAM3762 (N-terminal Hyp1-His overexpression construct). D. Schematic representation of pAM3763 (full length Hyp1-His overexpression construct). E. Schematic representation of pAM3760 (Truncated Hyp1-YFP:overexpression construct). F. Schematic representation of pAM3766 (full length Hyp1-YFP:overexpression construct)

CHAPTER IV

SUMMARY AND CONCLUSIONS

I chose to work on Hyp1 because of its several interesting features: (1). Overexpressing Hyp1 will cause an abnormal circadian phenotype, which shows that it must play some role in the circadian system. (2). The predicted transmembrane domain at the N-terminal half and cytochrome C-like structure at the C-terminal make it fit into our CikA functional model very well. Probably, this protein is CikA's interaction partner that docks CikA on the poles and ensures CikA's proper function. Moreover, the cytochrome-C homolog at the C-terminus makes us connect Hyp1 with cellular metabolism. More and more evidence in our lab shows that resetting the clock in *S. elongatus* is metabolism-dependent. It is likely that Hyp1 is the molecule that senses and controls redox status of the cell and therefore bridges CikA and cellular metabolism. Clarifying the function of Hyp1 will provide us with useful information for understanding how the signal is transmitted from the environment to the circadian system as well as how it is transmitted from the input pathway to the central oscillator. (3). It is interesting to notice that Hyp1 is not a universally conserved protein in cyanobacteria. Its homologs are only found in *Synechocystis* sp. PCC 6803, which indicates that cyanobacteria are likely to have diverse input pathways for their circadian systems.

REFERENCES

1. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997) *Protein Eng* **10**, 1-6.
2. Dunlap, J. C. (1996) *Annu Rev Genet* **30**, 579-601.
3. Ditty, J. L., Williams, S. B. & Golden, S. S. (2003) *Annu Rev Genet* **37**, 513-543.
4. Bell-Pedersen, D., Cassone, V. M., Earnest, D. J., Golden, S. S., Hardin, P. E., Thomas, T. L. & Zoran, M. J. (2005) *Nat Rev Genet* **6**, 544-556.
5. Dunlap, J. C., Loros J.J, DeCoursey P.J (2004) in *Chronobiology: Biological Timekeeping*, ed Dunlap, J. C (Sinauer Sunderland, MA), pp 53-64
6. Pittendrigh, C. S (1981) in *Handbook of Behavioral Neurobiology: Biological Rhythms*. ed J. Aschoff, (Plenum Press, New York.) pp. 57-80.
7. Golden, S.S., and L.A. Sherman. (1984). *J.Bacteriol.* **158**:36-42.
8. Kondo, T., Strayer, C. A., Kulkarni, R. D., Taylor, W., Ishiura, M., Golden, S. S. & Johnson, C. H. (1993) *Proc Natl Acad Sci U S A* **90**, 5672-5676.
9. Kondo, T. & Ishiura, M. (1994) *J Bacteriol* **176**, 1881-1885.
10. Ishiura, M., Kutsuna, S., Aoki, S., Iwasaki, H., Andersson, C. R., Tanabe, A., Golden, S. S., Johnson, C. H. & Kondo, T. (1998) *Science* **281**, 1519-1523.
11. Iwasaki, H., Nishiwaki, T., Kitayama, Y., Nakajima, M. & Kondo, T. (2002) *Proc Natl Acad Sci U S A* **99**, 15788-15793.
12. Hayashi, F., Suzuki, H., Iwase, R., Uzumaki, T., Miyake, A., Shen, J. R., Imada, K., Furukawa, Y., Yonekura, K., Namba, K. & Ishiura, M. (2003) *Genes Cells* **8**, 287-296.
13. Mori, T., Saveliev, S. V., Xu, Y., Stafford, W. F., Cox, M. M., Inman, R. B. & Johnson, C. H. (2002) *Proc Natl Acad Sci U S A* **99**, 17203-17208.
14. Xu, Y., Mori, T., Pattanayek, R., Pattanayek, S., Egli, M. & Johnson, C. H. (2004) *Proc Natl Acad Sci U S A* **101**, 13933-13938.
15. Pattanayek, R., Wang, J., Mori, T., Xu, Y., Johnson, C. H. & Egli, M. (2004) *Mol Cell* **15**, 375-388.
16. Xu, Y., Mori, T. & Johnson, C. H. (2000) *EMBO J* **19**, 3349-3357.

17. Kitayama, Y., Iwasaki, H., Nishiwaki, T. & Kondo, T. (2003) *EMBO J* **22**, 2127-2134.
18. Nishiwaki, T., Satomi, Y., Nakajima, M., Lee, C., Kiyohara, R., Kageyama, H., Kitayama, Y., Temamoto, M., Yamaguchi, A., Hijikata, A., Go, M., Iwasaki, H., Takao, T. & Kondo, T. (2004) *Proc Natl Acad Sci U S A* **101**, 13927-13932.
19. Lakin-Thomas, P. L. (2006) *J Biol Rhythms* **21**, 83-92.
20. Kutsuna, S., Nakahira, Y., Katayama, M., Ishiura, M. & Kondo, T. (2005) *Mol Microbiol* **57**, 1474-1484.
21. Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T. & Kondo, T. (2005) *Science* **308**, 414-415.
22. Tomita, J., Nakajima, M., Kondo, T. & Iwasaki, H. (2005) *Science* **307**, 251-4.
23. Uzumaki, T., Fujita, M., Nakatsu, T., Hayashi, F., Shibata, H., Itoh, N., Kato, H. & Ishiura, M. (2004) *Nat Struct Mol Biol* **11**, 623-631.
24. Williams, S. B., Vakonakis, I., Golden, S. S. & LiWang, A. C. (2002) *Proc Natl Acad Sci U S A* **99**, 15357-15362.
25. Hitomi, K., Oyama, T., Han, S., Arvai, A. S. & Getzoff, E. D. (2005) *J Biol Chem* **280**, 19127-19135.
26. Iwasaki, H., Williams, S. B., Kitayama, Y., Ishiura, M., Golden, S. S. & Kondo, T. (2000) *Cell* **101**, 223-233.
27. Smith, R. M. & Williams, S. B. (2006) *Proc Natl Acad Sci U S A* **103**, 8564-8569.
28. Takai, N., Nakajima, M., Oyama, T., Kito, R., Sugita, C., Sugita, M., Kondo, T. & Iwasaki, H. (2006) *Proc Natl Acad Sci U S A* **103**, 12109-12114.
29. Ivleva, N. B., Bramlett, M. R., Lindahl, P. A. & Golden, S. S. (2005) *EMBO J* **24**, 1202-1210.
30. Katayama, M., Kondo, T., Xiong, J. & Golden, S. S. (2003) *J Bacteriol* **185**, 1415-1422.
31. Miyagishima, S. Y., Wolk, C. P. & Osteryoung, K. W. (2005) *Mol Microbiol* **56**, 126-143.
32. Mutsuda, M., Michel, K. P., Zhang, X., Montgomery, B. L. & Golden, S. S. (2003) *J Biol Chem* **278**, 19102-19110.
33. Zhang, X., Dong, G. & Golden, S. S. (2006) *Mol Microbiol* **60**, 658-668.

34. Ivleva, N. B., Gao, T., LiWang, A. C. & Golden, S. S. (2006) *Proc Natl Acad Sci U S A* **103**, 17468-17473.
35. Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000) *Annu Rev Biochem* **69**, 183-215.
36. Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S. K., Zhu, Y., Ishima, R., Liu, D., Tong, K. I., Kurokawa, H., Qian, H., Inouye, M. & Ikura, M. (1999) *Nat Struct Biol* **6**, 729-734.
37. Kageyama, H., Kondo, T. & Iwasaki, H. (2003) *J Biol Chem* **278**, 2388-2395.
38. Phizicky, E. M. & Fields, S. (1995) *Microbiol Rev* **59**, 94-123.
39. Sambrook, J., and Russell, D.W., (1989). *Molecular Cloning*, ed Argentine J (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)
40. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res* **25**, 3389-3402.
41. Clerico, E. M., Ditty, J.L, and Golden S.S. (2006) in *Methods in Molecular Biology*, eds Rosata E (Humana, Totowa, NJ), pp 155-171.

APPENDIX A

A SUMMARY OF USING A YEAST TWO-HYBRID SYSTEM TO IDENTIFY PROTEINS THAT INTERACT WITH KAI B

KaiB is the oscillator protein for which we have the least information regarding function. We don't get any clues about KaiB function from bioinformatics analysis. If we could find some interacting partners for KaiB, it will provide us some clues about KaiB's function. The yeast two-hybrid assay is the most popular system used to investigate protein-protein interactions. I used the Clontech MATCHMAKER two-hybrid system 3 to do a yeast two-hybrid screen with KaiB as a bait. In this assay, interaction between a bait protein and a prey protein reconstitutes a transcription factor and induces nutritional selection and colorimetric reporter genes.

I fused the *kaiB* gene from *S. elongatus* to the GAL4 DNA-binding domain in pGADT7. The bait construct was used to transform the yeast strain AH109 and the expression of KaiB in yeast was confirmed with western blot.

We have a prey library obtained from Dr. H. Iwasaki (Nagoya University) that contains genomic DNA fragments from an *S. elongatus* PCC 7942 *kaiABC* null strain cloned in pGKDT7. We have only a very small amount of prey library *E. coli* cells left; therefore, I amplified the prey library by plating and extracted the plasmids from the library clones according to the protocol in the MATCHMAKER GAL4 Two-Hybrid System 3 & Libraries User Manual. I introduced the prey library into yeast strain AH109 which already contained a KaiB bait and gradually increased the stringency of selective requirements to select positive yeast clones (Figure 16).

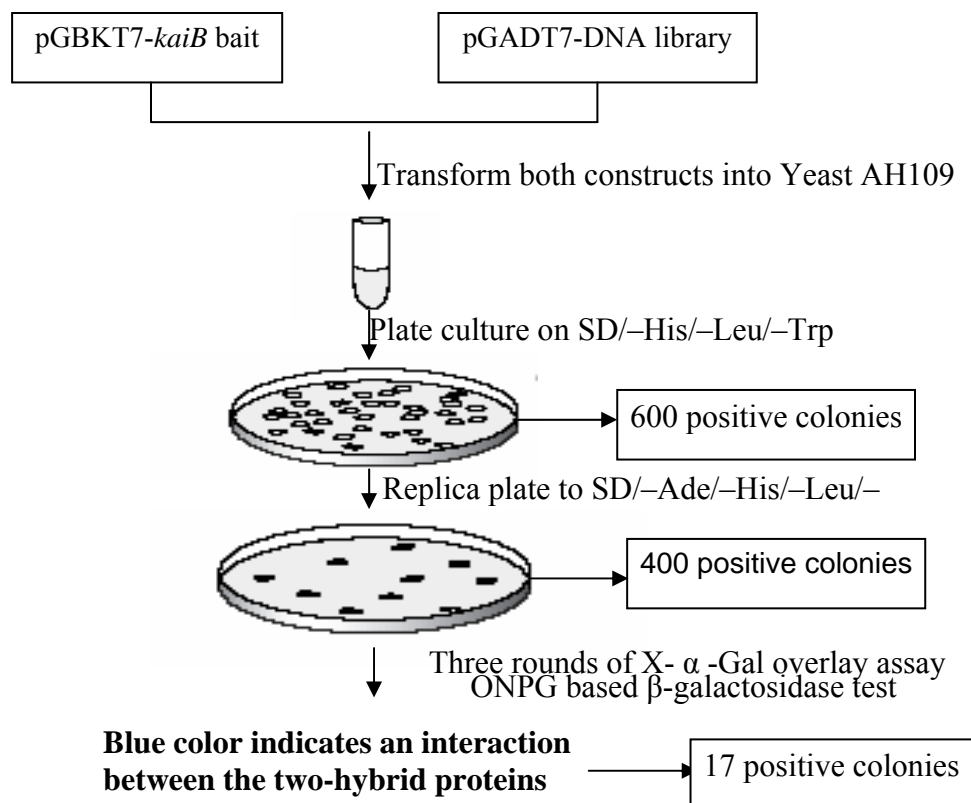


Fig. 16. Flow chart of yeast two-hybrid assay. pGBKT7-*kaiB* and pGADT7-DNA library were introduced into yeast AH 109 sequentially. The yeast transformants were selected on SD/-His/-Leu/-Trp. The positive clones were transferred to higher stringency plates SD/-Ade/-His/-Leu/-Trp. Three rounds of X- α -Gal overlay assay were carried out to narrow down the potential positive clones. Those that turned blue for at least two times were further examined with an ONPG-based β -galactosidase test. Eventually, 17 positive colonies qualifying for the highest stringency requirement were obtained.

Eventually, I got 17 positive yeast colonies qualifying for the highest stringency requirement. I numbered each positive yeast clone in the low stringency screen from 1 to 600. After three rounds of screening, 17 yeast colonies (#208, #215, #218, #221, #222, #225, #226, #230, #247, #253, #266, #407, #410, #478, #481, #494, #533) were found to qualify through all the selective stringencies.

Clones #208, #215, #218, #221, #222, #225, #230, #247, #266, #407, and #494 produced blue color for three times when performing X- α -Gal overlay assay on filters. Clones #226, #253, #410, #478, #532, and #533 demonstrated blue color twice when

performing X- α -Gal overlay assay on filters and they also produced a color change in the ONPG based β -galactosidase test.

I extracted prey plasmids from final positive clones and used them to transform *E. coli* DH10B cells. The yeast strain can be transformed by two or more plasmids simultaneously. Thus, the plasmid DNA isolated from each positive yeast colony will be a mixture of the DNA-BD/bait plasmid and at least one type of AD/prey library plasmid. I digested five plasmids isolated from randomly picked *E. coli* clones that derived from each yeast positive clone and compared their restriction maps. All the plasmids that have different restriction maps were sent for sequencing (for example, #208-1 and #208-4 are two different clones with different restriction maps). The genes and proteins they encode were identified for each clone through bioinformatics analysis (Table 3).

Table 3: The sequencing and retest results of 17 positive clones

| Colonies | Protein | Retest |
|--|--|---------------|
| #225-1, #266-3, #230-1, #247-2, #533-1 | Arginine biosynthesis bifunctional protein ArgJ. | Self-Activate |
| #208-1, #221-2, #222-1, #226-1, #494-4, #532-1, #253-1, #481-1, #494-5 | Bacterioferritin comigratory protein | Self-Activate |
| #407-1 | Predicted transcriptional regulator | Self-Activate |
| #208-4 | Hypothetical protein syc2327_d | Self-Activate |
| #494-1 | ATP phosphoribosyltransferase | Self-Activate |
| #221-3 | Hypothetical protein Selo03000035 | Self-Activate |
| #215-2 | pANL53 | No colony |
| #218-1, #410-1 | DnaK(HSP70-2) | Self-Activate |
| #478-1 | Hypothetical protein syc1607_d | Self-Activate |

The sequencing results showed that there are two frequently appearing positive clones. Clones #225-1, #266-3, #230-1, #247-2, and #533-1 are different positive clones that encode the same protein: arginine biosynthesis bifunctional protein ArgJ. ArgJ

catalyzes two activities that are involved in the cyclic pathway of arginine biosynthesis: the synthesis of acetylglutamate from glutamate and acetyl-CoA, and of ornithine by transacetylation between acetylornithine and glutamate. Clones #208-1, #221-2, #222-1, #226-1, #494-4, #532-1, #253-1, #481-1, and #494-5 encode two proteins: cell division protein Ftn2 homolog and bacterioferritin comigratory protein (BCP). In the genome, they are neighbors to each other but are transcribed in opposite directions (Figure 17). Sequence analysis showed that, in the prey vector, BCP will be translated and cell division protein Ftn2 homolog will not. Therefore, it should be BCP that interacts with KaiB. In *E. coli*, BCP is a putative bacterial member of the TSA/AhpC family, and it was characterized as a thiol peroxidase. BCP showed a thioredoxin-dependent thiol peroxidase activity. BCP preferentially reduces linoleic acid hydroperoxide rather than H₂O₂ and t-butyl hydroperoxide with the use of thioredoxin as an *in vivo* immediate electron donor.

It is interesting to note that there are three BCP homolog genes in *S. elongatus* PCC 7942. One of them is located upstream of CikA, but it is not the one we found in the yeast two-hybrid screen.

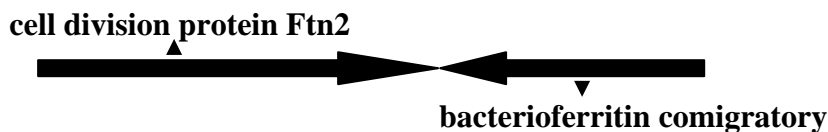


Fig. 17. The gene structure of the bacterioferritin comigratory protein locus. In the *S. elongatus* PCC 7942 genome, a gene encoding a cell division protein Ftn2 homolog is located upstream of a bacterioferritin comigratory protein (BCP) encoding gene. They are in opposite orientation.

Since #494-4, #221-2, and #208-1 encode bacterioferritin comigratory protein, and they appeared most frequently in the screening, it is likely that #208-4, #221-3, and #494-1 are not real prey proteins, but rather just entered the yeast together with plasmids that encode bacterioferritin comigratory protein.

There are some commonly encountered false positives in yeast two-hybrid assays using genes from various organisms. These include elongation factors, ferritin, lamin, cytoskeletal proteins, ribosomal RNA clones, and ubiquitin. Clones #218-1 and #410-1

encode the heat shock protein (HSP) DnaK (HSP70-2). HSPs are the most common proteins found to be false positives. Clone #407-1 encodes a predicted transcriptional regulator. Transcriptional regulators may activate the transcription of the marker gene without any interaction with KaiB. Thus, all of the clones may be false positives.

Clones #215-2 and #410-1 are from the *S. elongatus* plasmid pANL53. Clone #478-1 encodes a hypothetical protein. We do not have any functional clues for either of them.

I retested the interactions to test whether the clones are false positives. When the prey protein constructs and pGADT7 (empty bait vector without *kaiB*) were cotransformed into yeast AH109, even without KaiB, 21 of the prey proteins could activate transcription of the reporter genes: the yeast colonies grew on SD/-Ade/-His/-Leu/-Trp and in the X- α -Gal overlay assay, they turned blue (Table 3).

When plasmids #215-2 were cotransformed with pGADT7-KaiB into yeast AH109, no colonies showed up, even in low selective stringency plates. Thus, all the positive clones we obtained are false positives.

It is interesting to note that all of plasmids encoding arginine biosynthesis bifunctional protein ArgJ or bacterioferritin comigratory protein were clones of identical sequences rather than independent clones of the ORFs. The same prey library was used previously by Dr. Shannon R. Mackey to find CikA's interaction partner, and these common false positives did not show up in her screen. She used the original prey library, but I amplified the prey library. Most likely, in the prey library amplification, some false positive plasmids were selectively amplified; thus, they showed up again and again in my screening, which interferes the screen. In the future, we can reconstruct a new library to do yeast two-hybrid screens.

VITA

Name: Haitao Guo

Address: Dr. Susan. S. Golden
BSBW 314C BSBE, Mail Stop 3258
Biology Department
Texas A&M University
College Station, TX, 77843

Email address: haitao.guo@gmail.com

Education: B.S., Biochemistry, Central China Normal University, 2000
M.S., Genetics, Institute of Hydrobiology, Chinese Academy of Sciences,
2003
M.S., Microbiology, Texas A&M University, 2007